



Departamento de Bioquímica y Biología Molecular
Facultad de Biología
Universidad de Santiago de Compostela

El minisatélite Msh43: una sorprendente secuencia del genoma humano

Paula Barros Fernández

Tesis Doctoral

Marzo de 2008

Jaime Gómez Márquez, Catedrático de Bioquímica y Biología Molecular de la Universidad de Santiago de Compostela, y **Francisco Boán Fernández**, Doctor en Biología e Investigador del Departamento de Bioquímica y Biología Molecular de la Universidad de Santiago de Compostela, certifican:

Que la presente Tesis Doctoral, con el título: “*El minisatélite MsH43: una sorprendente secuencia del genoma humano*”, presentada por la Licenciada en Biología **Dña. Paula Barros Fernández**, ha sido realizada bajo nuestra dirección en el Departamento de Bioquímica y Biología Molecular de la Facultad de Biología de la Universidad de Santiago de Compostela y que, hallándose concluida, autorizamos su presentación para que pueda ser juzgada por el tribunal correspondiente.

Y para que así conste a los efectos oportunos, firmamos el presente certificado en Santiago de Compostela, a 14 de Marzo de 2008.

Dr. Jaime Gómez Márquez

Dr. Francisco Boán Fernández

A mi familia

Agradecimientos

En el momento de presentar este trabajo me gustaría expresar mi agradecimiento a todas aquellas personas que me han ayudado, tanto a nivel profesional como personal.

En primer lugar, debo agradecer a mis Directores, el Dr. Jaime Gómez y el Dr. Francisco Boán, la posibilidad de realizar esta investigación, así como el estar siempre disponibles para darme su consejo y apoyo.

No sería justo olvidar a la siempre paciente Dra. María Luz González, quien me inició en el mundo de la investigación.

También quiero mencionar a los otros seres que pueblan este extraño mundo que es el Departamento de Bioquímica y quienes han hecho que el trabajo fuese más llevadero. Miguel, Ester, Nora, Carmen, Fran, Amparo, Carla, Ana Canda, Mónica, Jorge,....

También quiero darle las gracias a mi familia porque siempre he contado y sé que siempre contaré con su apoyo.

Por último, tengo que dar las gracias a mi otra familia por estar siempre dispuestos a escucharme, aconsejarme, hacerme reír y disculpar que no estuviese cuando debía estar. Nieves, Berta, Esther, Oso, Bea, Pepo, Abraham, Juanito, Bea Orosa, Chukito, Xisela, Sebas, Solís, Pablo, Elías, Marta, Beña, Óskar,.....(y perdón a los que quedan sin nombrar)

Debo además, agradecer a la Dra. María Delclaux el habernos facilitado muestras de sangre de diversos primates.

Esta tesis ha sido financiada por el Ministerio de Educación y Ciencia (BFU2005-01083 y BFU2006-06708) y por la Xunta de Galicia (PGIDIT07PXIB200109PR).

Resumen

El minisatélite MsH43 es una secuencia del genoma humano cuyo estudio ha sido el principal objetivo de esta Tesis Doctoral. El análisis del polimorfismo de este locus nos permitió identificar ocho alelos. Las amplificaciones del locus MsH43 de individuos heterocigotos generan estructuras heterodúplex que les confiere un patrón electroforético específico que empleamos para identificar nuevos genotipos. La mayoría de los alelos aislados presentan la capacidad de formar estructuras G-cuadrúplex. Estas estructuras, estabilizadas por potasio, pueden interferir en la síntesis de DNA durante la PCR. Además, hemos analizado si la presencia de los G-cuadrúplex influye en la capacidad recombinogénica de MsH43, empleando para ello un sistema de recombinación homóloga *in vitro* diseñado en nuestro laboratorio. Los resultados de este análisis no detectaron diferencias en la capacidad recombinogénica entre un alelo que forma G-cuadrúplex y otro que no lo hace. Sin embargo, estos ensayos mostraron que el aumento de la fuerza iónica produce una mayor fidelidad en los eventos recombinatorios, ya que se incrementa la proporción de moléculas recombinantes que mantienen la secuencia original. Por otro lado, el análisis de la historia evolutiva de MsH43 reveló que esta secuencia se originó principalmente por procesos de *slippage*, mediante los cuales se fue expandiendo hasta alcanzar una organización de secuencia repetitiva en tándem. Según nuestra hipótesis, esta organización surgió durante la génesis de los Monos del Viejo Mundo. El análisis comparativo de las secuencias de MsH43 procedentes de distintas especies de primates, mostró que orangután y humano presentan el mayor grado de homología, lo que contrasta con la filogenia estándar. Para explicar esta paradoja sugerimos la existencia de un alelo ancestral, que denominamos alelo errante, que se ha mantenido prácticamente inalterado en el período evolutivo comprendido entre orangután y humano.

Summary

The MSH43 minisatellite is a sequence of the human genome whose study has been the main objective of the present Doctoral Thesis. The analysis of the MSH43 polymorphism allowed us to identify eight alleles. The amplifications of the MSH43 locus from heterozygous individuals generated heteroduplex structures that were employed to identify new genotypes by electrophoresis. The majority of the identified alleles present the ability to form G-quadruplex. This structure is stabilized by potassium and interferes with the DNA synthesis during PCR. Moreover, we have analyzed if the presence of the G-quadruplex could influence the recombinogenic capacity of MSH43, using for this purpose an in vitro homologous recombination system designed in our laboratory. The results of this analysis showed no differences in the recombinogenic capacity between a G-quadruplex forming allele and another one without such ability. Nevertheless, these assays showed that the increase of the ionic strength produces a higher fidelity in the recombinatory events, since the proportion of recombinant molecules that maintain the original sequence is increased. On the other hand, the evolutionary history of MSH43 revealed that this sequence was originated mainly by slippage events that provoked the sequence expansion until it reached a tandem sequence organization. According to our hypothesis, this organization arose during the genesis of Old World Monkeys. The comparative analysis of MSH43 sequences from different primate species showed that the orang-utan and human species presented the highest degree of homology which disagrees with the standard phylogeny. To explain this paradox we suggest the existence of an ancestral allele, the wandering allele, which was maintained almost unaltered in the evolutionary period comprised between orang-utan and human species.

Abreviaturas

BSA	Seroalbúmina bovina
CEPH	Centro para el estudio del Polimorfismo Humano
DSB	Corte de doble cadena
DMS	Dimetilsulfato
DTT	Ditiotreitol
LB	Medio de Luria-Bertani
Msh43	Minisatélite Msh43
PMSF	Fluoruro de fenilmetil sulfonilo
MNM	Monos del Nuevo Mundo
MVM	Monos del Viejo Mundo

Índice

Introducción	15
1. Secuencias repetitivas de DNA	17
2. El DNA microsatélite y minisatélite	19
2.1. El DNA minisatélite	20
2.2. El DNA microsatélite	25
2.3. Origen y evolución de los microsatélites y minisatélites	27
3. Análisis de secuencias repetitivas en tándem: artefactos durante la PCR	29
4. Los minisatélites Msh42 y Msh43	32
4.1. El minisatélite Msh42	32
4.2. El minisatélite Msh43	34
Objetivos	39
Materiales y métodos	43
1. Metodología general de clonación	45
2. Obtención de DNA genómico	45
3. Amplificaciones por PCR	46
4. Secuenciación de DNA	48
5. Análisis de estructuras G-cuadrúplex	48
6. Ensayo de interferencia de la síntesis de DNA	50
7. Ensayo de heterodúplex	51
8. Ensayos de recombinación homóloga intramolecular <i>in vitro</i>	51
8.1. Construcción de los sustratos plasmídicos	51
8.2. Preparación de extractos proteicos nucleares	54

8.3 Ensayo de recombinación homóloga intramolecular	55
Resultados y discusión	57
Análisis del polimorfismo de Msh43	59
Amplificación preferencial y G-cuadrúplex	65
Capacidad recombinogénica y G-cuadrúplex	77
Nacimiento e historia evolutiva de Msh43	95
Conclusiones	109
Bibliografía	113
Apéndice	127

INTRODUCCIÓN

La secuenciación del genoma humano (International Human Genome Sequencing Consortium, 2001, 2004; Venter et al., 2001) ha permitido conocer que aproximadamente el número de genes codificadores de proteínas se sitúa en torno a los 20000-25000. Sin embargo, los genes solamente constituyen ~2% del genoma humano. El resto del genoma está constituido por otro tipo de secuencias, la mayoría de las cuales es de naturaleza repetitiva. ¿Qué son las secuencias repetitivas? ¿Cómo es su estructura? ¿Cuál es su función? ¿Cómo se originaron? ¿Cómo evolucionan? En esta Tesis Doctoral abordaremos el estudio de una secuencia repetitiva del genoma humano con estructura de micro/minisatélite que denominamos MsH43.

1. SECUENCIAS REPETITIVAS DE DNA

La existencia de secuencias repetitivas en el genoma fue propuesta a partir de los resultados obtenidos en los experimentos de cinética de reasociación de cadenas de DNA desnaturalizado (Britten & Kohne, 1968). Así, se caracterizó al DNA repetitivo por su mayor tasa de renaturalización respecto al DNA no repetitivo o de copia única (Lewin, 2008). Posteriormente, se realizaron diversos estudios para conocer la estructura y localización concreta de este tipo de secuencias. El hecho de que no tuviesen ninguna función conocida hizo que algunos investigadores las consideraran DNA basura, un lastre genético, o “DNA egoísta”, remanente de genomas virales u otro tipo de secuencias que se comportan de forma parásita (Orgel & Crick, 1980). Sin embargo, posteriores investigaciones sobre distintas clases de DNA repetitivo han revelado que muchas de estas secuencias están implicadas en importantes procesos biológicos (Lewin, 2008).

En la Figura 1 se muestra una clasificación de los distintos tipos de secuencias repetitivas presentes en el genoma (Brown, 1999), siendo el principal criterio la disposición de las unidades repetitivas que las constituyen. Utilizando este criterio, las secuencias repetitivas se pueden dividir en dos grupos: dispersas, cuando las unidades que se repiten se encuentran distribuidas a lo largo del genoma, y en tándem, en las que las repeticiones se localizan agrupadas una tras otra. Por su parte, las secuencias repetitivas dispersas generalmente se clasifican en base a su mecanismo de transposición. Así, podemos encontrar las siguientes clases:

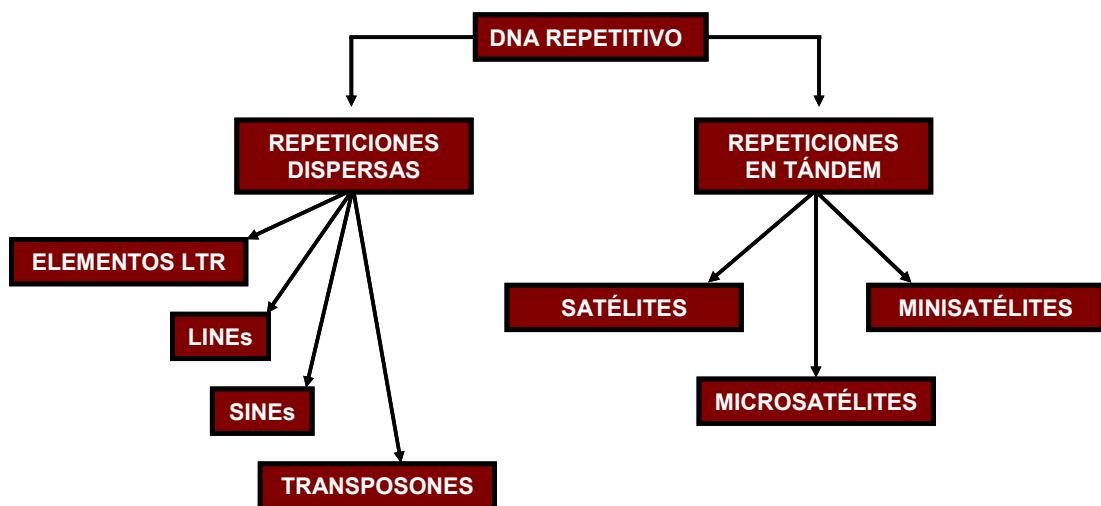


Figura 1. Clasificación del DNA repetitivo según su disposición y estructura en el genoma.

- Elementos LTR (long terminal repeats): transposones de RNA caracterizados por la presencia de largas repeticiones terminales en sus extremos.
- LINEs (long interspersed nuclear elements): elementos nucleares dispersos largos que poseen un gen tipo transcriptasa inversa probablemente involucrado en su transposición.
- SINEs (short interspersed nuclear elements): elementos nucleares dispersos cortos que carecen de un gen que codifique una transcriptasa inversa propia y que utilizan la transcriptasa inversa que producen otros retroelementos para su retrotranscripción. El ejemplo más representativo de este tipo de secuencias lo constituye la familia *Alu*.
- Transposones de DNA: secuencias transponibles que utilizan una transposasa para realizar una interacción directa DNA/DNA entre el transposón donante y el lugar aceptor.

En cuanto a las secuencias formadas por unidades repetitivas dispuestas en tándem, se pueden distinguir tres tipos de DNA atendiendo al tamaño de la unidad que se repite:

- Satélite: integrado por unidades repetitivas de un tamaño superior a 100 pb.
- Minisatélite: constituido por unidades repetitivas de 5/6-100 pb.
- Microsatélite: secuencias compuestas por repeticiones en tándem de 1-5/6 pb.

2. EL DNA MICROSATÉLITE Y MINISATÉLITE

Como se ha mencionado anteriormente, los microsatélites y minisatélites representan un tipo de secuencias repetitivas en tándem con una unidad que se repite de pequeño o mediano tamaño. El límite entre lo que se puede considerar una secuencia microsatélite o minisatélite es arbitrario, como lo demuestra la gran variedad en el rango del tamaño de la unidad repetitiva usado por los distintos investigadores para clasificarlas. Esto provoca que existan secuencias que se encuentran en la frontera entre micro- y minisatélite, es decir, que podrían ser consideradas tanto microsatélites como minisatélites. Como veremos, éste es el caso de la secuencia Msh43, objeto de estudio del presente trabajo. A esta clase de secuencias nos referiremos como micro/minisatélites. Teniendo presente este hecho, a continuación profundizaremos en algunos de los aspectos de la biología de estas secuencias repetitivas que quedan reflejados en la Figura 2.

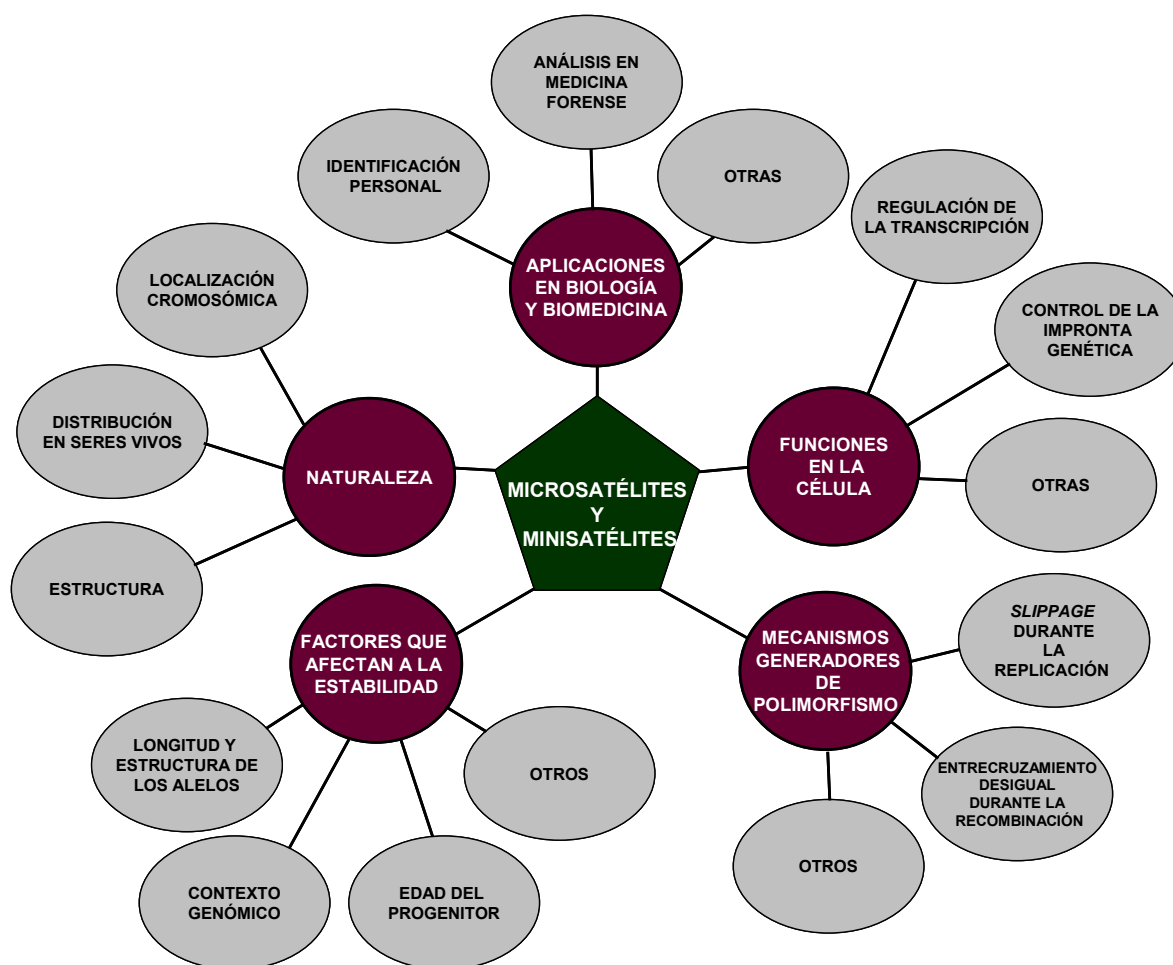


Figura 2. Diagrama que representa los principales aspectos del estudio de los microsatélites y los minisatélites.

2.1. EL DNA MINISATÉLITE

El término minisatélite fue empleado por primera vez en 1985 para referirse a secuencias constituidas por la repetición en tándem de bloques cortos de DNA, ricos en GC, muy polimórficos y que presentan homología con la secuencia *Chi* (Jeffreys et al., 1985a). Actualmente, el término minisatélite abarca a un grupo más amplio de secuencias repetidas en tándem, con un tamaño de la unidad que se repite entre 5/6-100 pb (Vergnaud & Denoeud, 2000). Su grado de polimorfismo es variable, existiendo secuencias monomórficas (Ali & Wallace, 1988), poco polimórficas (Boán et al., 1997) y muy polimórficas (Vogt, 1990; Bois & Jeffreys, 1999; Vergnaud & Denoeud, 2000). Este polimorfismo es frecuentemente debido a variaciones en el número de las unidades repetitivas que los componen, por lo que también se empleó el término VNTR (Variable Number of Tandem Repeats) para referirse a estas secuencias (Nakamura et al., 1987). La alta variabilidad que presentan algunos minisatélites los convierte en valiosos marcadores moleculares, siendo empleados en el desarrollo de técnicas de DNA *fingerprinting* para estudios de identificación personal (Jeffreys et al., 1985b), así como análisis en medicina forense (Jobling & Gill, 2004) y para la realización de mapas genéticos de ligamiento (Nakamura et al., 1987).

Los minisatélites están presentes en el genoma de prácticamente la totalidad de organismos, tanto procariotas como eucariotas: bacterias (Mrazek et al., 2007), levaduras (Andersen & Nillson-Tillgren, 1997), hongos (Giraud et al., 1998), plantas (Dallas, 1988) y animales (Jeffreys & Morton, 1987). En el genoma humano, aunque existen secuencias minisatélite distribuidas a lo largo de las distintas zonas de los diferentes cromosomas, su presencia es mayor en las regiones subteloméricas (Royle et al., 1988).

¿Qué funciones puede desempeñar el DNA minisatélite en la célula? Dependiendo del tipo de minisatélite y de su localización en el genoma estas secuencias pueden tener distintas funciones (Vergnaud & Denoeud, 2000). Así, la demostración de que algunos minisatélites son reconocidos específicamente por proteínas celulares (Collick & Jeffreys, 1990; Wahls et al., 1991; Wahls & Smith, 1994; Boán et al., 1997; Fukuda et al., 2001) apoya dicha funcionalidad. Sin embargo, también hay minisatélites que probablemente no tengan ningún papel en la biología del genoma y que se encuadran dentro del DNA que algunos investigadores denominan como DNA basura o egoísta (Orgel & Crick, 1980). A continuación, veamos algunos ejemplos que apuntan a una función biológica de este tipo de secuencias.

Existen minisatélites que se encuentran localizados en el entorno de determinados genes y que pueden influir en su transcripción. En este sentido, se ha podido detectar la asociación entre determinados alelos de secuencias minisatélite y un aumento del riesgo de padecer diabetes mellitus insulino dependiente (Bennet et al., 1995) o epilepsia mioclónica progresiva tipo 1 (Laloti et al., 1997). Además, se ha observado que expansiones o contracciones en el número de unidades repetitivas de DNA minisatélite, situado en zonas codificantes, provocan la aparición o desaparición de aminoácidos en la cadena polipeptídica y, por lo tanto, la aparición de distintas variantes proteínicas. Ejemplos de este tipo son los genes de las apolipoproteínas (Boguski et al., 1986) y el receptor D4DR de la dopamina (Van Tol et al., 1992; Lichter et al., 1993) donde el polimorfismo de los minisatélites asociados a estos genes se ha relacionado con la generación de distintas variantes alélicas que muestran cambios en su afinidad por la unión a lípidos y a la dopamina respectivamente. Otro ejemplo de la posible influencia de un minisatélite en un proceso biológico lo constituye la observación de que algunos minisatélites localizados en los intrones de algunos genes afectan al *splicing* (Turri et al., 1995). Por otro lado, se conoce la existencia de DNA minisatélite asociado a lugares frágiles en los cromosomas (Sutherland et al., 1998), asimismo, se han descrito minisatélites que pueden estar implicados en el control de la impronta genética de algunos loci (Chaillet et al., 1995; Neumann et al., 1995). Por último, se ha demostrado la relación de estas secuencias con procesos recombinatorios, pudiendo potenciar la recombinación homóloga *in vitro* (Walhs et al., 1990; Boán et al., 1998; Walhs & Moore, 1998), lo que podría indicar que este tipo de secuencias pueden representar “hotspots” de recombinación, es decir, lugares donde la tasa de recombinación es superior a la media del resto del genoma (Walsh, 1998).

El polimorfismo de las secuencias minisatélite puede deberse a variaciones en el número y/o en el orden de las unidades repetitivas (Jeffreys et al., 1990). Atendiendo a la organización que presentan, se pueden distinguir cinco clases de minisatélites: puro, interrumpido puro, compuesto, interrumpido compuesto y complejo (Tabla 1). Los minisatélites puros son aquellos que están constituidos por la repetición de un único motivo; cuando en esta secuencia aparece un pequeño motivo de copia única que corta la repetición en tándem, al minisatélite se le denomina interrumpido puro. Los minisatélites compuestos son los que están formados por repeticiones de dos motivos; al igual que en el caso anterior, cuando en esta secuencia existe un pequeño fragmento de copia única que corta la repetición en tándem del motivo, al minisatélite se le denomina

interrumpido compuesto. Finalmente, cuando la secuencia minisatélite consta de más de dos unidades repetitivas que pueden estar interrumpidas por secuencias cortas de copia única se les denomina minisatélite complejo.

Tabla 1. Clasificación de los minisatélites según su organización (modificado de Buschiazzi & Gemmel, 2006)

Clase	URD*	Ejemplo
Puro	1	-(TAGGG) ₁₂ -
Puro interrumpido	1	-(TAGGG) ₈ -(TGAGG)-(TAGGG) ₃ -
Compuesto	2	-(TAGGG) ₉ -(TGAAC) ₅ -
Compuesto interrumpido	2	-(TAGGG) ₉ -(TAAGG)-(TGAAC) ₄ -
Complejo	>2	-(TAGGG) ₄ -(TGAAG) ₇ -CAATG-(TTGGG) ₃ -

*Unidades repetitivas diferentes

Los procesos que generan la variabilidad en los minisatélites son fundamentalmente dos: el entrecruzamiento desigual y el deslizamiento de cadena durante la replicación (*slippage*). El entrecruzamiento desigual, que fue propuesto para explicar la evolución de una secuencia repetitiva en tándem (Smith, 1976; Jeffreys et al, 1998), se produce durante el proceso de recombinación y debido a la similitud entre las unidades repetitivas. Así, las moléculas de DNA se alinean a veces de forma errónea, dando lugar a duplicaciones o deleciones que conducen a un alargamiento o acortamiento de la secuencia repetitiva (Figura 3). Por su parte, el *slippage* tiene lugar

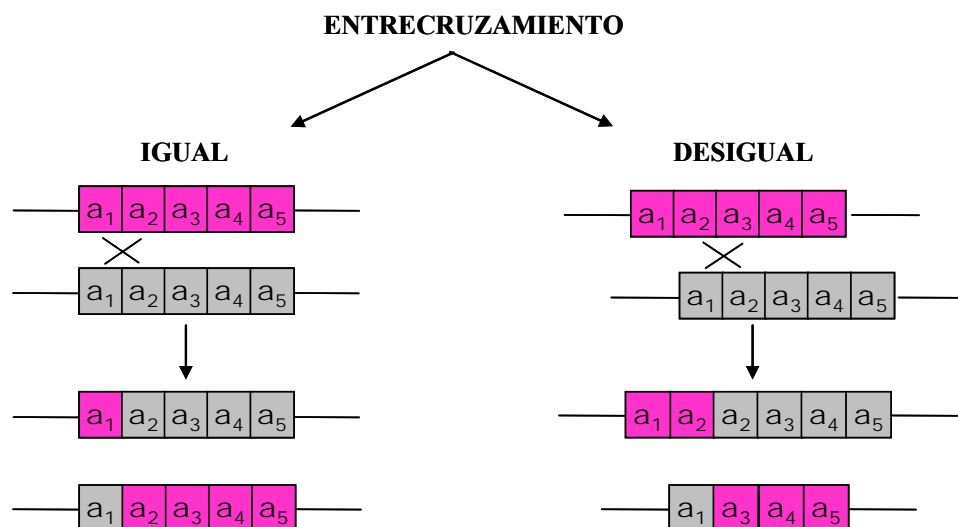


Figura 3. Esquema del proceso de entrecruzamiento durante la recombinación. A la izquierda se muestra las secuencias que se generan tras un entrecruzamiento igual y a la derecha, las producidas tras un entrecruzamiento desigual. Los cuadrados denotados con letras representan las unidades repetitivas y las “X” los lugares de entrecruzamiento y posterior recombinación.

cuando durante la replicación las hebras de DNA se alinean de forma no homóloga provocando la formación de bucles de cadena sencilla. Si la hebra que se desliza es la que sirve de molde se producirá una delección, mientras que si es la de nueva síntesis, tendrá lugar una duplicación (Figura 4). Este proceso ha sido propuesto como el principal determinante en la generación y evolución de DNA repetitivo en tándem (Levinson & Gutman, 1987; Buschiazzo & Gemmel, 2006). Además de estos mecanismos se han descrito otros que pueden dar lugar a polimorfismos como la conversión génica, en la que se produce una transferencia no recíproca de DNA de una molécula a su homóloga (Jeffreys et al., 1994) o la hibridación de cadena sencilla, en la que tras producirse un corte de doble cadena en la secuencia repetitiva, los extremos se reseccionan dando lugar a dos cadenas sencillas complementarias que pueden hibridar de forma no homóloga (Pâques & Haber, 1999).

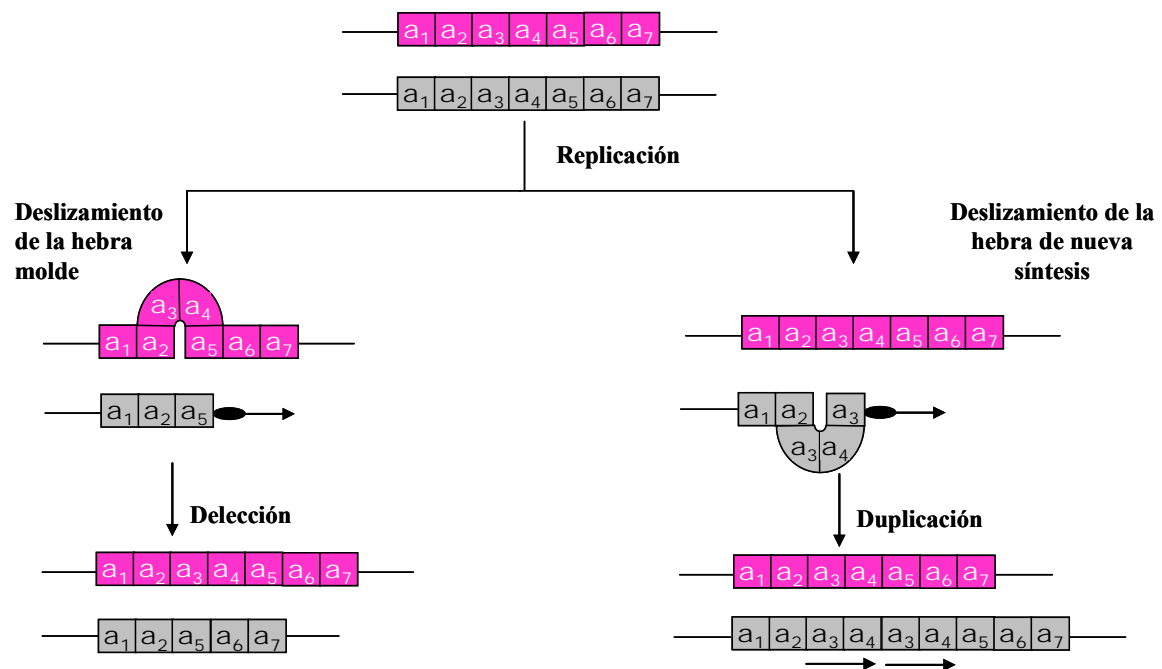


Figura 4. *Slippage* como generador de variabilidad en una secuencia repetida en tándem. Los cuadrados denotados con letras representan las unidades repetitivas y las elipses con la flecha representan la maquinaria replicativa y el sentido de la síntesis de DNA.

La hipótesis de que algunos minisatélites presentes en los genomas eucariotas pudieran tener la función de ser *hotspots* de recombinación, surgió como consecuencia de dos observaciones: el alto grado de polimorfismo que presentan algunas de estas secuencias y la similitud de la secuencia consenso de las unidades repetitivas de muchos minisatélites con la secuencia *Chi* (Jeffreys et al., 1985a), conocido punto caliente de

recombinación en procariotas (Stahl et al., 1975). A favor de esta hipótesis que relaciona minisatélites y recombinación está el hecho de que estas secuencias se sitúan preferentemente en regiones subteloméricas de los cromosomas (Royle et al., 1988), zonas se caracterizan por altas tasas de recombinación y además en que algunos minisatélites han sido localizados en los quiasmas cromosómicos (Chandley & Mitchell, 1988), estructuras que reflejan a nivel citológico los lugares de entrecruzamiento meiótico. Por otra parte, se ha descrito que la tasa de mutación para algunos loci minisatélite humanos es diferente entre mujeres y hombres (Buard & Vergnaud, 1994; Henke & Henke, 1995), lo que recuerda a las diferencias observadas en los patrones específicos de sexo en la recombinación meiótica (Hunt & Hassold, 2002). Otra importante evidencia de la implicación de los minisatélites en recombinación, fue la demostración de que el minisatélite humano CEB1 constituye un *hotspot* meiótico, estimulando la existencia de entrecruzamientos y de conversiones génicas complejas (Buard et al., 2000) y que una secuencia minisatélite consenso sintética, estimula hasta 13,5 veces la recombinación homóloga entre dos plásmidos, sin capacidad para replicarse, introducidos en la línea celular humana EJ (Wahls et al., 1990). Esta capacidad de una secuencia minisatélite para potenciar la recombinación homóloga *in vitro* también se puso de manifiesto en los experimentos llevados a cabo en nuestro laboratorio (Boán et al., 1998), en los que establecía una comparación de la capacidad recombinogénica del minisatélite Msh42 con una secuencia no repetitiva del mismo tamaño.

Es conocida la existencia de inestabilidad de algunas secuencias minisatélite en línea germinal (Bois & Jeffreys, 1999), lo cual ha conducido a la realización de diversos estudios para analizar los factores que afectan a dicha inestabilidad. Algunos de estos trabajos se llevaron a cabo empleando técnicas como la SP-PCR y MVR-PCR. A continuación, señalamos algunas de las conclusiones que se han extraído de dichos estudios. Se ha observado la existencia de minisatélites que presentan una polaridad para la aparición de mutaciones, bien en su zona 5' (Jeffreys et al., 1994) ó 3' (Armour et al., 1993), pudiendo estar relacionada con la presencia de *hotspots* adyacentes al minisatélite que podrían provocar su inestabilidad (Jeffreys et al., 1998). En este sentido, se ha observado la existencia de polimorfismos nucleotídicos en el DNA flanqueante 5' del minisatélite MS32 que tienen relación directa con la tasa de mutación detectada en este locus como consecuencia de procesos recombinatorios (Monckton et al., 1994). Otra conclusión fue que el tamaño y la homogeneidad de los alelos de ciertos

minisatélites son factores que influyen en el grado de inestabilidad de un minisatélite, sobre todo en aquella que es generada por reorganizaciones intralélicas (Buard et al., 1998). También se ha mostrado que cuanto mayor es el grado de divergencia entre los alelos parentales, mayor es la tasa de generación de mutantes que presentan variaciones en el tamaño de la secuencia minisatélite (He et al., 2002; Jauert & Kirkpatrick, 2005) y que existe una correlación entre la edad del padre en el momento de la concepción y la tasa de mutación del DNA minisatélite (Andreassen et al., 2002). Otro dato destacable es que, en algunos minisatélites, las mutaciones que dan lugar a nuevos alelos con variaciones en el tamaño ocurren en una tasa mayor en la línea germinal que en células somáticas (Cederberg & Rannug, 2006). Finalmente, en algunos de estos estudios se han descrito que algunos alelos mutantes parecen originarse mediante eventos recombinatorios más complejos que el *slippage* o el entrecruzamiento desigual. Para explicar estas reorganizaciones se han propuesto diversos modelos que combinan *slippage* y recombinación, iniciados por cortes del DNA dentro del minisatélite (Bois & Jeffreys, 1999) o en su entorno (He et al., 2002).

2.2 EL DNA MICROSATÉLITE

Como se mencionó anteriormente, el término microsatélite define un conjunto de secuencias repetitivas en tándem, cuya unidad de repetición es de un tamaño comprendido entre 1-5/6 pb. Al igual que los minisatélites, los microsatélites se pueden clasificar en base a la estructura de secuencia repetitiva en 5 clases: puro, interrumpido puro, compuesto, interrumpido compuesto y complejo (Buschiazzo & Gemmel, 2006) (Tabla 2).

Tabla 2. Clasificación de los microsatélites.

Clase	Número de unidades repetitivas diferentes	Ejemplo
Puro	1	-(CA) ₁₂ -
Puro interrumpido	1	-(CA) ₈ -(CT)-(CA) ₃ -
Compuesto	2	-(CA) ₉ -(GAA) ₅ -
Compuesto interrumpido	2	-(CA) ₉ -(CAA)-(GAA) ₄ -
Complejo	>2	-(CA) ₄ -(T) ₇ -CAAT-(CTT) ₃ -

La presencia de secuencias microsatélite ha sido demostrada en prácticamente todos los organismos eucarióticos y procarióticos. Así, por ejemplo, se ha descrito en

bacterias (Mrazek et al., 2007), levaduras (Toth et al., 2000), hongos (Mouhamadou et al., 2007), plantas (Moncada et al., 2006), animales (Beckmann & Weber, 1992) y humanos (Toth et al., 2000).

En cuanto a su distribución cromosómica, el DNA microsatélite se encuentra disperso a lo largo de los cromosomas, tanto en zonas codificadoras como no codificadoras. Sin embargo, en el caso de algunas secuencias microsatélite parece que dicha distribución no se produce al azar, ya que algunos loci que presentan determinadas unidades repetitivas presentan localizaciones específicas (Toth et al., 2000). Un ejemplo de esto lo encontramos en los microsatélites compuestos de tripletes predominan en regiones codificadoras (Borstnik & Pumpernik, 2002). La causa de este sesgo en la distribución no se conoce, aunque se piensa que podría estar relacionada con la implicación de algunos microsatélites en determinados procesos biológicos.

Se han descrito diversas evidencias experimentales que sugieren la implicación de algunos microsatélites en la regulación de la actividad génica. Un ejemplo de esto lo constituye la observación de que la delección de parte de un microsatélite presente en el promotor del gen *c-KI-ras* produce un marcado descenso en la tasa de transcripción de dicho gen (Hoffman et al., 1990). En algunos casos, el número de unidades repetitivas que componen el DNA microsatélite es el factor clave en la regulación de la actividad génica, existiendo genes que sólo se expresan en presencia de un número determinado de unidades repetitivas del microsatélite. Por ejemplo, la presencia del oligonucleótido (GAA)₁₂ en el promotor del gen *LacZ* de *E. coli* permite su expresión, mientras que su sustitución por los oligonucleótidos (GAA)₁₄₋₁₆ o por (GAA)₅₋₁₁ no permiten que dicho gen sea expresado (Liu et al., 2000). También se ha demostrado que secuencias microsatélite localizadas en intrones pueden afectar a la transcripción de los genes en los que se encuentran, como sucede en el caso del gen de la tirosin hidroxilasa, en el que la presencia de un microsatélite en el primer intrón actúa como un elemento regulador de su tasa de transcripción (Meloni et al., 1998). También se ha propuesto la relación entre DNA microsatélite y cáncer y, en esta dirección, se ha sugerido que la acumulación de inestabilidad en microsatélites puede contribuir al desarrollo de tumores malignos como el cáncer gástrico (Leung et al., 2000). Por último, los microsatélites pueden funcionar como puntos calientes de recombinación (Templeton et al., 2000), idea que ha sido apoyada por experimentos realizados con levadura (Trecó & Arnheim, 1986), humanos (Majewski & Ott, 2000) y células de mamíferos (Wahls & Moore, 1990).

El DNA microsatélite presenta un elevado grado de polimorfismo, hecho que comparte con muchos minisatélites, y que al igual que en el caso de éstos ha servido para que sean utilizados como marcadores en diversos estudios de genética de poblaciones (Jarne & Lagoda, 1996), en análisis forenses (Tamaki & Jeffreys, 2005) y en reconstrucciones filogenéticas (Harr et al., 1998). Los procesos que generan polimorfismo en las secuencias microsatélite, son esencialmente los mismos que citamos anteriormente como principales responsables de la variabilidad del DNA minisatélite, es decir, el entrecruzamiento desigual durante la recombinación y el *slippage*. Sin embargo, se ha sugerido que en el caso de los microsatélites el *slippage* es el principal proceso generador de polimorfismo (Tautz & Schlötterer, 1994; Buschiazzo & Gemmel, 2006).

Diferentes estudios han puesto de manifiesto la influencia de diversos factores sobre la inestabilidad del DNA microsatélite. En este sentido, se ha observado que las tasas de mutación pueden variar significativamente entre distintos loci microsatélite y/o alelos del mismo locus dependiendo de su estructura. Así, los alelos más largos son a menudo más polimórficos que los cortos (Ellegren., 2000), lo que sugiere o bien un aumento de la inestabilidad de la maquinaria replicativa cuando se enfrenta a secuencias repetitivas largas (Wierdl et al., 1997) o simplemente, que los procesos de *slippage* son más probables en secuencias repetitivas de mayor tamaño (Buschiazzo & Gemmel, 2006). Otra característica que parece alterar la tasa de mutación en el DNA microsatélite es la longitud de la unidad repetitiva, de manera que la inestabilidad es inversamente proporcional al tamaño del motivo que se repite (Chakraborty et al., 1997). Por otro lado, se ha puesto de manifiesto que microsatélites que se encuentran próximos pueden influir unos en la evolución de los otros (Udupa et al., 2004). Otro factor a tener en cuenta es la composición de las secuencias que flanquean al DNA microsatélite. Así, se ha observado que el contenido en GC de estas secuencias o la proximidad de islas CpG (Brock et al., 1999) han mostrado su influencia en la inestabilidad de DNA microsatélite. Por otro lado, existen evidencias que correlacionan la tasa de mutación de los microsatélites de un individuo con la edad del padre en el momento de la concepción (Gusmao et al., 2005; Brinkmann et al., 1998).

2.3. ORIGEN Y EVOLUCIÓN DE LOS MICROSATÉLITES Y MINISATÉLITES

El origen de las secuencias micro- y minisatélite continúa siendo un tema de debate. Actualmente, los estudios realizados sobre la génesis de dichas secuencias

repetitivas derivan en el planteamiento de dos hipótesis (Buschiazzo & Gemmel, 2006). La primera hipótesis sugiere que el DNA micro- y minisatélite se originaría a partir de secuencias de copia única de las cuales surgirían unidades repetitivas como consecuencia de inserciones (Dieringer & Schlötterer, 2003), sustituciones (Messier et al., 1996) o ambas (Boán et al., 2004). La segunda hipótesis plantea que el origen de estas secuencias repetitivas en un determinado locus puede deberse a un elemento móvil que contendría en su secuencia lugares proclives a la formación de este tipo de secuencias repetitivas (Wilder & Hollocher, 2001). En este contexto, sería posible que el elemento móvil contuviese secuencias repetitivas en tándem ya organizadas con lo que la transposición de estos elementos daría lugar a la dispersión de los microsatélites y minisatélites en el genoma. Sea cual sea su origen, una vez que se han generado las duplicaciones en tándem, a partir de una secuencia de copia única o de elementos transponibles, se produciría la expansión de este protomicrosatélite o protominisatélite mediante procesos de *slippage* principalmente (Rose & Falush, 1998). ¿Es necesario un número mínimo determinado de unidades repetitivas para que se produzca la expansión inicial? Mientras algunos investigadores cuestionan la existencia de ese número (Pupko & Graur, 1999; Perez et al., 2005), otros plantean que la cantidad de unidades repetitivas necesarias se situaría entre 4-8 (Rose & Falush, 1998; Sibly et al., 2001; Shinde et al., 2003; Lai & Sun, 2003). Es también probable que algunos loci presenten dicho número crítico mientras que en otros no exista.

Una vez se ha producido la expansión inicial, dando lugar a una organización del DNA que se corresponde con una secuencia repetitiva en tándem, esta secuencia sufrirá variaciones en su longitud, por la pérdida y ganancia de unidades repetitivas, de nuevo mediante dos procesos, *slippage* durante la replicación y/o entrecruzamiento desigual durante la recombinación. El *slippage* podría generar de forma mayoritaria la adición o pérdida de una sola unidad repetitiva en cada evento (Weber & Wong, 1993; Amos et al., 1996; Brinkmann et al., 1998) o la de varias unidades (Huang et al., 2002), mientras que los procesos recombinatorios podrían producir expansiones o contracciones de mayor tamaño (Richard & Pâques, 2000). A pesar de que no existen pruebas irrefutables, se cree que el proceso de *slippage* es el principal generador de variabilidad tanto para microsatélites (Tautz & Schlötterer, 1994) como para minisatélites (Levinson & Gutman, 1987). Las diferencias entre las tasas de contracción y expansión generadas por este proceso parecen depender del tamaño de la secuencia repetitiva. Así, en humanos, mientras la tasa de expansión es lineal y constante, la tasa de contracción es

inicialmente baja para después aumentar exponencialmente, a medida que el microsatélite aumenta de tamaño (Xu et al., 2000). Esta predisposición hacia la contracción de los microsatélites más largos puede ser el resultado de la acción de fuerzas selectivas que hacen que los micro- y minisatélites, y por lo tanto los genomas, mantengan un rango de tamaño razonable. Otros eventos que contribuirían a la generación de polimorfismo de estas secuencias serían las sustituciones, así como las pequeñas inserciones y deleciones que pueden producir interrupciones en el patrón repetitivo de las mismas.

¿Cómo seguirían evolucionando las secuencias micro/minisatélites una vez se ha producido su origen y crecimiento? Se ha propuesto que el crecimiento de la longitud del micro/minisatélite estaría limitado debido al equilibrio existente entre las expansiones y contracciones (Buschiazzi & Gemmel, 2006). Sin embargo, las interrupciones en la secuencia repetitiva provocadas por mutaciones puntuales se seguirían produciendo hasta que la acumulación de las mismas diese lugar a una secuencia de copia única, que incluiría pequeños fragmentos de la secuencia repetitiva original. Además, se podrían producir deleciones de gran tamaño, acelerando la degeneración de la secuencia repetitiva. Al resultado de este proceso se la ha denominado “muerte” del micro/minisatélite (Taylor et al., 1999). Estos procesos harían que la organización repetitiva estuviese tan deteriorada que su estructura se correspondería con la que presentaban las secuencias iniciales a partir de las cuales se generaron los micro/minisatélites. Por lo tanto, a partir de esa secuencia se podría volver a originar de nuevo un micro/minisatélite completándose así el “ciclo vital” de este tipo de secuencias repetitivas de DNA.

3. ANÁLISIS DE SECUENCIAS REPETIDAS EN TÁNDEM: ARTEFACTOS DURANTE LA PCR

Los microsatélites y minisatélites son valiosos marcadores moleculares, lo cual ha derivado en su amplia utilización en diversos ámbitos como el análisis de ligamiento (Nakamura et al., 1987), la identificación de individuos en medicina forense (Gill et al., 1985) y la determinación de relaciones de parentesco entre individuos (Jeffreys et al., 1985c). La principal técnica empleada para el análisis de estas secuencias es la PCR. Sin embargo, la peculiar estructura de este tipo de secuencias hace que esta técnica sea

susceptible de producir artefactos que pueden conducir a interpretaciones erróneas de los resultados.

Entre estos artefactos se encuentran los productos derivados de *slippage* (Hite et al., 1996), los heterodúplex (Anglani et al., 1990), la amplificación preferencial (Walsh et al., 1992) y la formación de estructuras por parte del DNA molde que interfieren en la polimerización, como los G-cuadrúplex (Woodford et al., 1994). El *slippage* se produce, como mencionamos anteriormente, cuando durante la amplificación, la polimerasa se desliza sobre la cadena molde generando, además del producto esperado, productos distintos que presentan duplicaciones o deleciones que pueden ser detectadas mediante un análisis electroforético (Figura 5). En cuanto a la formación de los heterodúplex, que son el resultado de la unión de dos cadenas procedentes de distintas moléculas de DNA, éstos se generan durante la renaturalización de los productos de amplificación, cuando las cadenas de un individuo heterocigoto, en lugar de hibridar cada una con su complementaria, lo hacen con la del otro alelo. Estos heterodúplex van a presentar bucles en las zonas no complementarias provocando una alteración en la migración electroforética que permitirá su detección (Figura 5). La amplificación preferencial consiste en la amplificación más favorable de uno de los alelos respecto del otro, en un individuo heterocigoto. Se han propuesto diversas causas para explicar este fenómeno: i) diferencias significativas en la longitud o en la riqueza en GC entre los alelos, ii) cantidades muy pequeñas de muestra que pueden producir variaciones estocásticas en el número de copias obtenidas de cada alelo, y iii) existencia de polimorfismos en la secuencia diana del cebador que provoquen que en uno de los alelos hibriden mal o incluso no hibriden. Finalmente, la aparición de artefactos durante la PCR también puede ser consecuencia de la adopción por parte del DNA molde de estructuras que interfieran en la polimerización. Un ejemplo de estas estructuras lo constituyen los G-cuadrúplex (Woodford et al., 1994), los cuales están formados por la asociación de cuatro hebras de DNA que presentan planos superpuestos de tétradas de guaninas. Estas tétradas son el resultado de la unión de cuatro guaninas mediante puentes de hidrógeno denominadas uniones de Hoogsteen. En el centro de la estructura y entre cada dos tétradas se sitúa un catión que estabiliza el G-cuadrúplex (Gellert et al., 1962), y que generalmente es potasio (Sen & Gilbert, 1990).

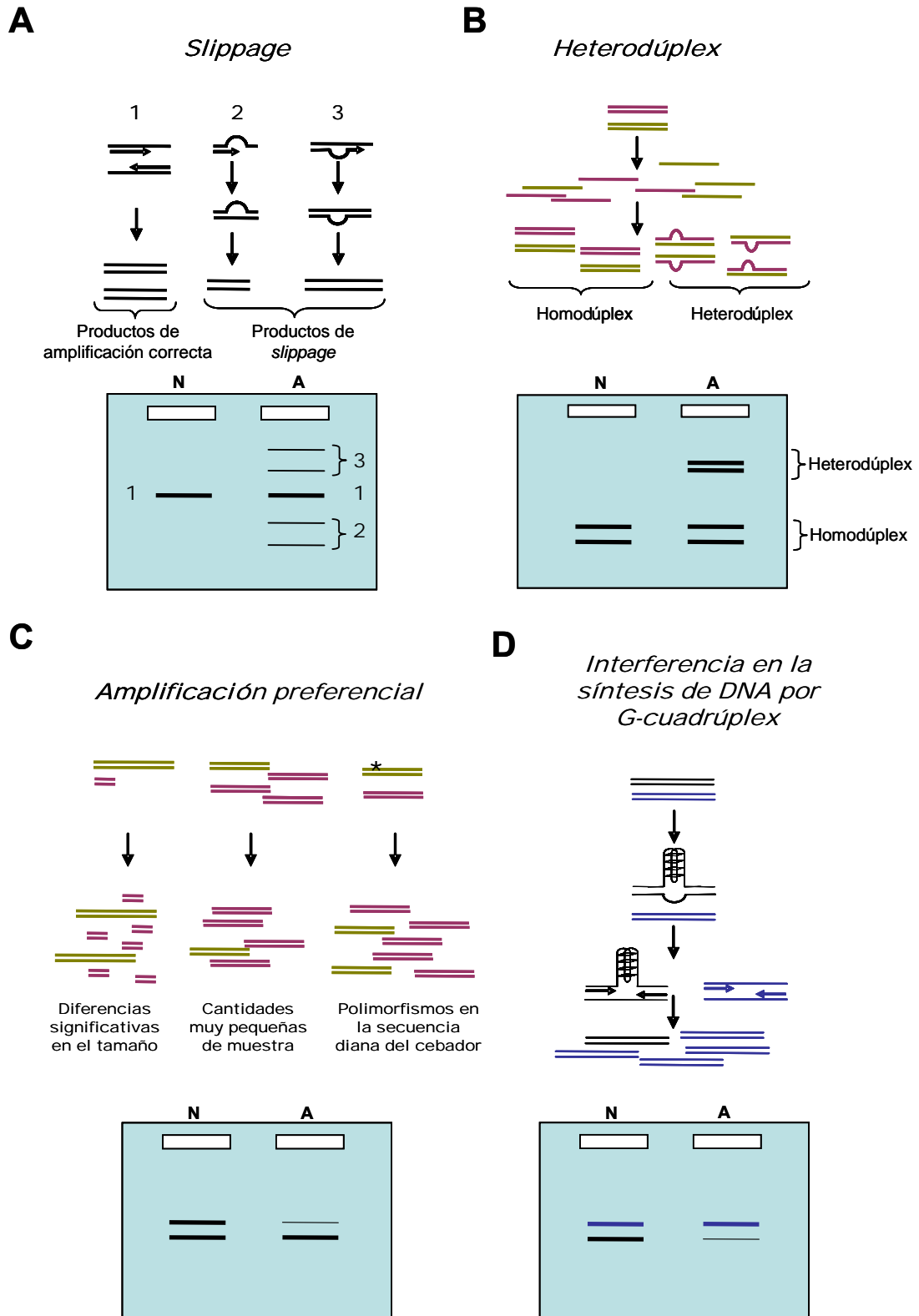


Figura 5. Representación esquemática de los mecanismos que pueden producir artefactos durante la PCR. En el esquema de todos los geles en el carril N se muestra lo que se obtendría en una amplificación normal y en el carril A lo que se produciría como consecuencia del artefacto correspondiente.

4. LOS MINISATÉLITES MsH42 Y MsH43

A principios de los años 80 se demostró experimentalmente la existencia de secuencias homólogas entre determinadas regiones del genoma de mamíferos y el de algunos miembros de la familia de los Herpesvirus (Puga et al., 1982; Peden et al., 1982, Heller et al., 1982). La hibridación de una genoteca de ratón con la sonda rica en GC *Sma*IA obtenida del virus Herpes simplex tipo1, permitió aislar varios clones que presentaban secuencias cortas repetidas en tándem con un alto contenido en GC (Gómez-Márquez et al., 1985). Posteriormente, con el mismo fragmento viral se sondeó una genoteca de hígado fetal humano (Lawn et al., 1978), lo que permitió aislar y caracterizar en nuestro laboratorio varios fagos recombinantes. El análisis de los insertos de estos clones mostró que la homología se debía a repeticiones en tándem ricas en GC (González, 1993), al igual que ocurría con los clones murinos mencionados anteriormente. Entre los clones aislados de esta genoteca humana se encuentran los denominados $\lambda 42$ y $\lambda 43$, conteniendo los minisatélites MsH42 y MsH43, respectivamente.

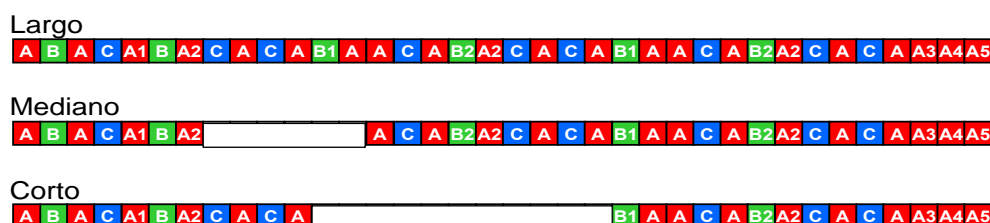
4.1. EL MINISATÉLITE MsH42

El análisis del clon $\lambda 42$ permitió caracterizar el minisatélite MsH42 (González, 1993). El minisatélite MsH42 fue el primero en ser estudiado en profundidad en nuestro laboratorio. Este minisatélite se encuentra en el intrón 13 del gen Q8WUJ3, localizado en la región subtelomérica del brazo largo del cromosoma 15 y presenta un bajo grado de polimorfismo. En la población humana se han caracterizado tres alelos denominados largo, mediano y corto, constituidos por 36, 30 y 25 unidades repetitivas de 9-10 pb, respectivamente (Boán et al., 1997) (Figura 6).

El análisis de su historia evolutiva (Boán et al., 2004) puso de manifiesto que el minisatélite MsH42 se originó a partir de una secuencia progenitora sin una estructura de minisatélite bien definida. Su nacimiento tuvo lugar, durante la evolución de los primates, antes de la divergencia entre los monos del Viejo y Nuevo Mundo. El minisatélite evolucionó probablemente a través de eventos como inserciones/delecciones, mutaciones sencillas y expansiones/contracciones mediadas por entrecruzamiento desigual durante la recombinación y/o *slippage* durante la replicación. Es interesante destacar que el polimorfismo en las poblaciones actuales de humanos, chimpancés y gorilas parece haberse generado a partir de un único alelo

ancestral, puesto que las secuencias del minisatélite en las tres especies presentan una gran homología entre ellas, tanto en secuencia como en organización de las repeticiones.

MINISATÉLITE Msh42



Código de las Repeticiones

A TGGGAGAGGC	A2 TGGGAAAGGC	A4 TGGGAAAGAC	B TGGGATTGC	B2 TGAGATTGC
A1 TGACAGAGGC	A3 TGGGAGAGAC	A5 TGGGAAAGA	B1 TGTGATTGC	C TGGGAGAGC

Figura 6. Estructura de los alelos del minisatélite Msh42 caracterizados en nuestro laboratorio. Las cajas blancas presentes en las secuencias de los alelos mediano y corto representan fragmentos delecionados con respecto al alelo largo.

Existen tres evidencias experimentales que relacionan el minisatélite Msh42 con procesos recombinatorios: i) en el entorno de este minisatélite se encuentran pequeñas secuencias asociadas a la recombinación, como son los motivos *Chi* y los oligómeros (TGTTTTTGG y CACATGT) implicados en la recombinación V(D)J en los genes de las inmunoglobulinas (Sakano et al., 1981), ii) la región Msh42 es reconocida específicamente por proteínas de extractos de testículo, hígado y cerebro de ratón (Boán et al., 1997; Rodríguez, 2002) y iii) Msh42 es capaz de sufrir procesos de *slippage* durante la PCR, dando lugar a la aparición de bucles de cadena sencilla que podrían estimular eventos recombinatorios (Boán et al., 1997). Estas evidencias llevaron a analizar la posible relación entre Msh42 y procesos recombinatorios. La conclusión de estos estudios fue que la región Msh42 es una secuencia que potencia la recombinación homóloga intramolecular *in vitro* cuando se compara con una secuencia control no repetitiva de tamaño similar (Boán et al., 1998, 2002). Posteriormente, el desarrollo de un ensayo *in vitro*, que se denominó ensayo DSB (corte de doble cadena), permitió detectar la generación de los cortes DSB que podrían iniciar los procesos recombinatorios, así como aislar y caracterizar los productos de recombinación

originados a partir de las moléculas cortadas. Uno de los resultados obtenidos fue la detección de dos patrones de corte diferentes dependiendo de la presencia o no de Mg^{2+} en el ensayo (Blanco et. al., 2005). Por otro lado, el análisis de los recombinantes reveló que el principal responsable de la formación de esas moléculas en el ensayo DSB era un mecanismo de tipo recombinación no homóloga por unión de extremos terminales (NHEJ) favorecido por microhomologías (Boán et al., 2006).

Los resultados obtenidos en nuestro ensayo DSB presentaban grandes diferencias con respecto a los mostrados por otros investigadores, a pesar de que las metodologías empleadas eran similares. Estas diferencias radicaban en que en nuestros experimentos el DNA era degradado rápidamente y la detección de multímeros lineales o moléculas circulares era muy baja; sin embargo, otros autores obtenían los resultados opuestos, escasa degradación y una elevada ligación con la consiguiente formación de multímeros. Después de diversos experimentos se concluyó que la predominancia de uno u otro evento depende de la relación DNA/proteína empleada. Cuando los valores de dicha relación son bajos se inhibe la actividad nucleásica, debido a que existen suficientes proteínas que se unen al DNA protegiéndolo de la degradación y potenciando la ligación, con la consiguiente generación de multímeros y moléculas circulares. Sin embargo, a medida que la relación DNA/proteína aumenta, la degradación se va haciendo mayor debido a que la cantidad de proteína no es suficiente para proteger al DNA de la actividad nucleásica (Blanco et al., 2004).

4.2. EL MINISATÉLITE Msh43

Otro de los clones aislados de la genoteca de hígado fetal humano empleando como sonda el fragmento *Sma*IA fue el denominado λ 43 (González, 1993). El análisis del fago recombinante λ 43, mediante digestión con enzimas de restricción y Southern blot reveló que la homología con el DNA viral se restringía a un fragmento de ~4,8 Kb, producto de la digestión con el enzima *Eco*RI. Este fragmento fue subclonado en pUC18 generando el plásmido pHH43 (Figura 7).

Para determinar la secuencia de DNA implicada en la hibridación, se digirió el plásmido pHH43 con *Eco*RI/*Pvu*II, dando lugar a un fragmento de ~1Kb que fue subclonado en el bacteriófago M13. La secuenciación de este fragmento, denominado

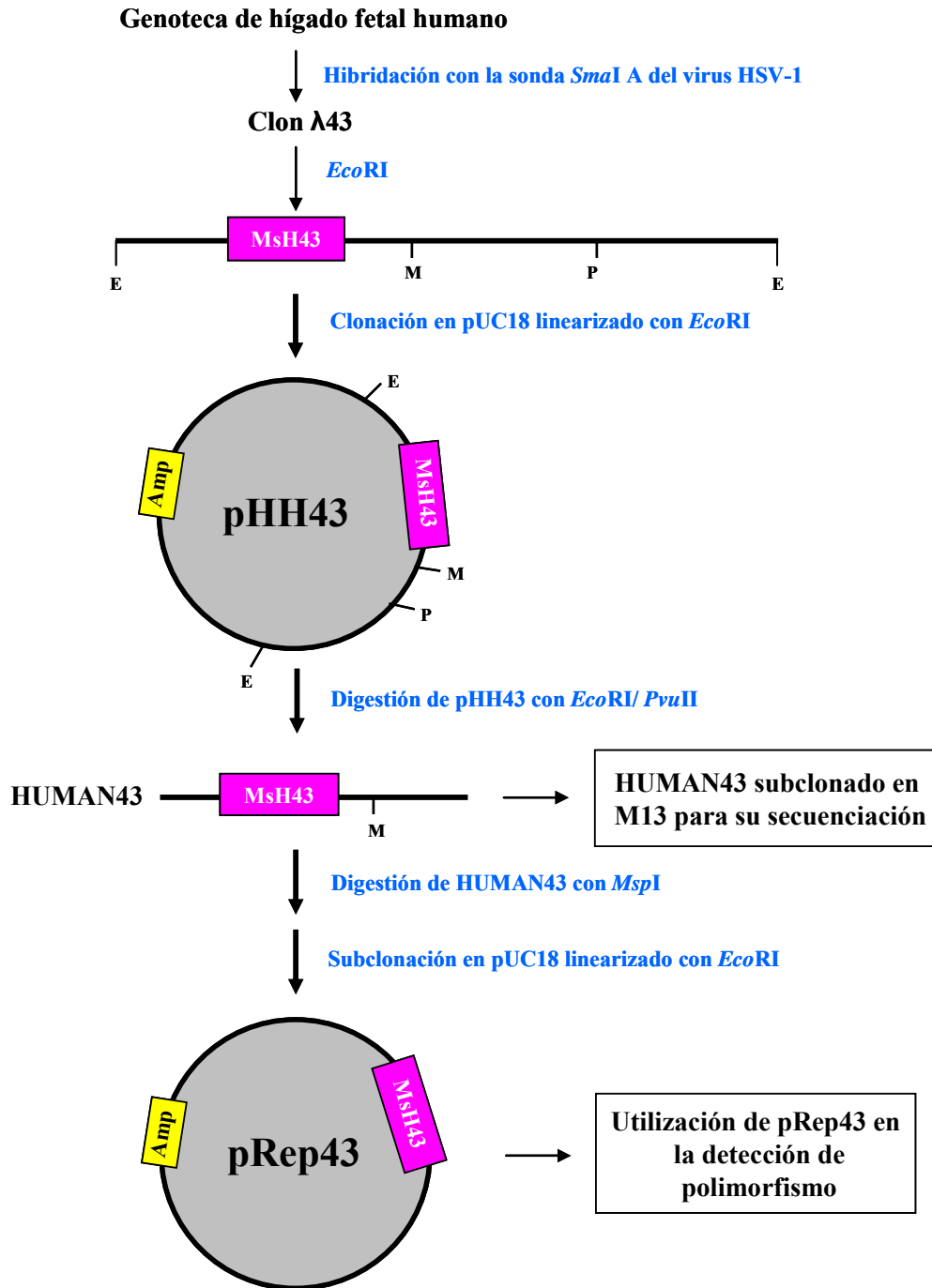


Figura 7. Aislamiento y subclonaje del minisatélite humano MsH43. La caja **MsH43** representa el minisatélite humano MsH43 y la caja **Amp** representa el gen de la resistencia a la ampicilina. **E**, *EcoRI*; **M**, *MspI*; **P**, *PvuII*.

HUMAN43, reveló que está constituido por 956 pb, dentro del se encuentra una secuencia repetida en tándem de 456 pb rica en GC situada entre los nucleótidos 36-492 (Figura 8). Dicha secuencia está constituida por repeticiones en tándem de motivos de

5-6 pb, lo que permitió incluirla dentro del grupo de los minisatélites y se le denominó MsH43. Cabe destacar además, la presencia de una secuencia *Alu*, entre los nucleótidos 820-956.

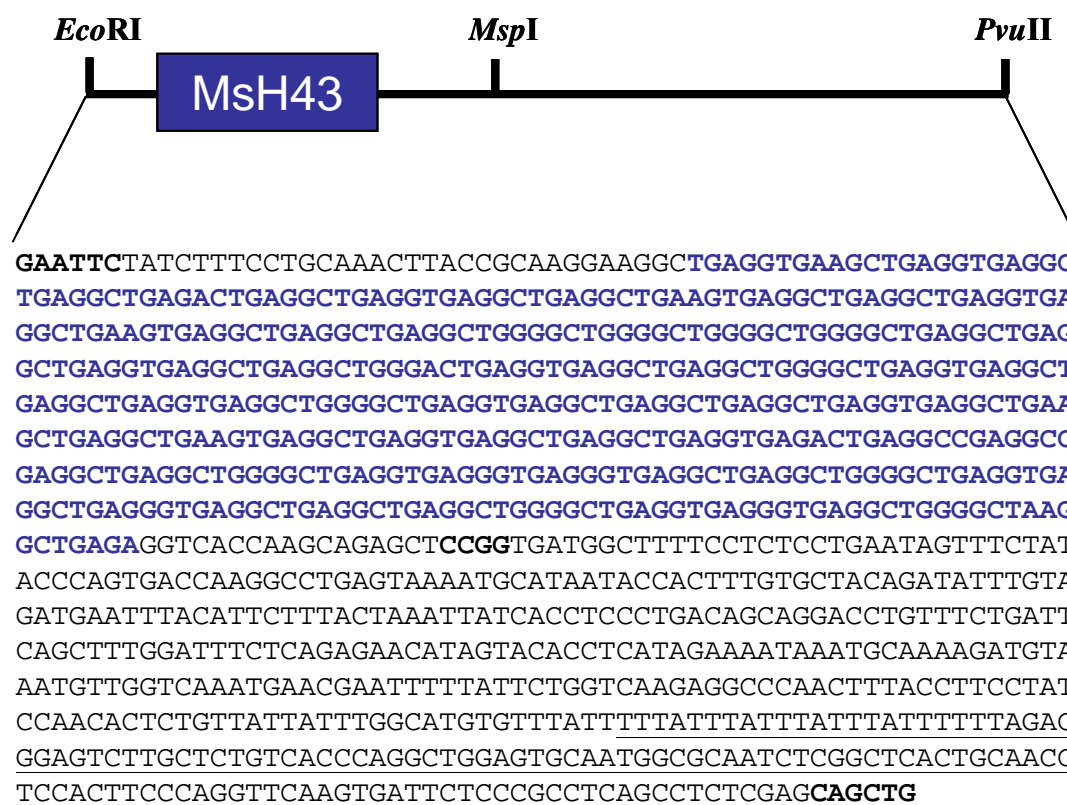


Figura 8. Secuencias del fragmento HUMAN43. La secuencia que corresponde al minisatélite MsH43 aparece en color azul. Las dianas de restricción de los enzimas *EcoRI*, *MspI* y *PvuII* se destacan en negrita. La secuencia subrayada corresponde a parte de una secuencia *Alu*.

Una vez determinada la secuencia de HUMAN43, se decidió abordar el análisis del polimorfismo del minisatélite MsH43. Para realizar este estudio fue necesario la obtención de una sonda específica con la que hibridar muestras genómicas humanas. Por ello se digirió la región HUMAN43 con *MspI*, para así eliminar la secuencia perteneciente a la familia *Alu*, ya que estas secuencias son muy abundantes en el genoma humano y podrían enmascarar los datos asociados específicamente al minisatélite. La digestión de la región HUMAN43 con *MspI* dio lugar a un fragmento de 510 pb que fue subclonado en un plásmido pUC18 generando el plásmido pRep43 (Figura 7). El estudio del polimorfismo se realizó mediante hibridaciones de pRep43

con muestras genómicas humanas digeridas con el enzima de restricción *Pst*I. El resultado de estos experimentos apuntaba a la existencia de distintas variantes alélicas en el locus MSH43 (González, 1993).

En esta Tesis Doctoral continuaremos con el estudio del minisatélite humano MSH43, analizando su polimorfismo, filogenia evolutiva y capacidad recombinogénica.

OBJETIVOS

Los objetivos que nos hemos propuesto durante la realización de esta Tesis Doctoral han sido:

- 1.- Analizar el polimorfismo del minisatélite MsH43 en la población humana.
- 2.- Estudiar el origen e historia evolutiva de MsH43 a través de su análisis en primates.
- 3.- Investigar la capacidad recombinogénica de MsH43 mediante un ensayo de recombinación homóloga intramolecular *in vitro*.

MATERIALES Y MÉTODOS

1. METODOLOGÍA GENERAL DE CLONACIÓN

La metodología general de clonación empleada en el presente trabajo de investigación se encuentra descrita en el manual de laboratorio de Sambrook & Russell (2001). Los protocolos para la preparación de DNA plasmídico a pequeña escala (miniprep), a gran escala (maxiprep) y para secuenciación se realizaron siguiendo métodos basados en la lisis alcalina. La preparación de células competentes y transformación de bacterias competentes *E. coli* DH5 α (recA⁻, endA⁻, ϕ 80 LacZ Δ M15) se llevó a cabo mediante el protocolo de Inoue et al. (1990).

2. OBTENCIÓN DE DNA GENÓMICO

Para el análisis del polimorfismo del minisatélite MsH43 se emplearon muestras de saliva de 132 individuos caucasoides sanos no emparentados. En el estudio filogenético se emplearon muestras de sangre de 11 especies de primates: 7 chimpancés (*Pan troglodytes*), 6 gorilas (*Gorilla gorilla*), 2 orangutanes (*Pongo pygmaeus*), 2 gibones (*Hylobates lar*), 3 macacos (2 *Macaca nigra* y 1 *Macaca tonkeana*), 1 mandril (*Mandrillus sphinx*), 2 mangabeyes (*Lophocebus albigena*), 1 tamarindo (*Saguinus oedipus*), 2 cebus (*Cebus apella*) y 1 lemur (*Varecia variegata*). En la Figura 9 se muestra una clasificación de los primates.

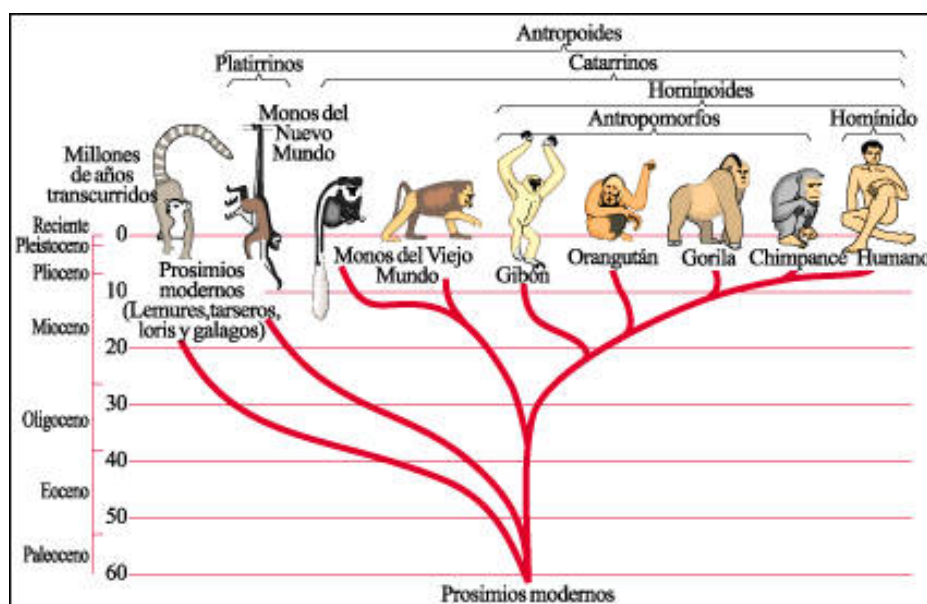


Figura 9. Árbol filogenético de primates

La mayoría de estas muestras fueron cedidas por María Delclaux del Zoológico de Madrid. Las diferentes muestras fueron procesadas para la extracción del DNA genómico empleando el QIAamp DNA blood kit (QIAGEN). También se emplearon muestras de DNA procedentes de la familia de referencia 1347 del Centro de Estudios del Polimorfismo Humano (CEPH) cedidas por el Dr. Xavier Estivill.

3. AMPLIFICACIONES POR PCR

Las secuencias de los oligonucleótidos empleados como cebadores en las distintas reacciones de PCR se muestran en la Tabla 3. Las reacciones para analizar el polimorfismo del locus Msh43 en humanos se realizaron en un volumen final de 25 µl conteniendo 50 ng de DNA genómico, 0,3 µM de cada uno de los cebadores P02.1 y P02.2, 0,2 mM dNTPs, 1,5 mM MgCl₂ y 0,5 U de Taq polimerasa con su correspondiente tampón. En la optimización de la amplificación del locus Msh43 se realizaron reacciones de PCR en las que se emplearon los tampones mostrados en la Tabla 4. También se llevaron a cabo reacciones empleando Taq polimerasa de distintos proveedores: Roche, Genecraft y Ecogen, con sus correspondientes tampones. Las amplificaciones se realizaron en un termociclador PTC-200 (MJ Research), bajo las siguientes condiciones: 95°C 1 min, 56°C 30 s y 72°C 40 s durante 29 ciclos seguidos de un ciclo adicional con una extensión de 5 min.

Tabla 3. Cebadores empleados en las amplificaciones.

Cebador	Secuencia (5'→3')
P02.1	GAATTCTATCTTTCCTGCAAAC
P02.2	GGCCTTGGTCACTGGGTATA
P1	CAGCTACAATTAGGCATTTATG
P2.1	GCATTTTACTCAGGCCTTGG
P2.2	AGGTAAAGTTGGGCCTTTTG
P2.3	GGCCTTTTGACCAGAATAAA
P3	CCAGAGAGCTCTTCAGGAG
P3.1	CTCTGNNNGTGTCTTCTGTGGG

Para la amplificación de DNA en primates, las reacciones se realizaron en un volumen total de 50 µl conteniendo 0,3 µM de cada cebador, 0,2 mM dNTPs, 1,5 mM MgCl₂ y 0,5 U de Taq polimerasa de Ecogen. Los cebadores empleados, así como las

condiciones de amplificación para cada especie fueron las siguientes: P1/P02.2, 95°C 1 min, 52°C 30 s y 72°C 40 s para chimpancés y gorilas; en el caso de los orangutanes y gibones las condiciones fueron las mismas excepto el tiempo de extensión que fue de 2,5 min; P1/P2.1, 95°C 1 min, 60°C 30 s y 72°C 3,5 min para los Monos del Viejo Mundo (macacos, mandriles y mangabeyes); P3/P2.2, 95°C 1 min, 58°C 30 s y 72°C 1 min para Monos del Nuevo Mundo (cebus y tamarindo) y P3.1/P2.3, 95°C 1 min, 53°C 30 s y 72°C 30 s para el lemur. Todas las reacciones se llevaron a cabo mediante 29 ciclos y un ciclo final con una extensión de 5 min.

En algunos casos, los diferentes productos de PCR que correspondían a alelos del minisatélite MsH43 fueron clonados en el plásmido pGEM-T Easy (Promega) para su posterior secuenciación. Este plásmido es un vector preparado para clonar productos de PCR generados por la Taq polimerasa, que añade una adenina en cada extremo 3' de los fragmentos de DNA producidos.

Tabla 4. Composición de los tampones empleados en la amplificación de MsH43.

Tampón (1X)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Tris-HCl (mM)	10	67	10	10	10	10	67	10	67	67	67	67	67	67
pH	8,3	8,3	8,8	8,3	8,3	8,3	8,8	8,8	8,3	8,8	8,8	8,8	8,8	8,8
KCl (mM)	50	50	50	-	50	50	-	-	-	50	-	50	-	12,5
(NH₄)₂SO₄ (mM)	-	-	-	16	-	16	16	16	16	-	16	16	-	-
Tween-20 (%)	-	-	-	-	0,01	-	0,01	0,01	0,01	0,01	-	0,01	0,01	0,01
Tampón (1X)	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Tris-HCl (mM)	67	67	67	67	67	67	67	67	67	67	67	67	67	67
pH	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8	8,8
KCl (mM)	25	37,5	62,5	75	87,5	100	42	36	30	24	18	12	6	-
(NH₄)₂SO₄ (mM)	-	-	-	-	-	-	2	4	6	8	10	12	14	24
Tween-20 (%)	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01

4. SECUENCIACIÓN DE DNA

Los fragmentos de DNA fueron secuenciados mediante un sistema de secuenciación automática cíclica utilizando el dGTP BigDye Terminator v3.0 Sequencing Kit (Applied Biosystems), según las especificaciones del fabricante, excepto en el hecho de que las reacciones se reescalaron para ser realizadas a la mitad del volumen final indicado (de 20 μ l a 10 μ l). El DNA plasmídico se preparó por el protocolo de maxiprep o de miniprep para secuenciación (Sambrook & Russell, 2001). Algunos productos de PCR se secuenciaron directamente cuando su tamaño era inferior a 700 pb. Para ello, los productos de PCR fueron tratados enzimáticamente con exonucleasa I y fosfatasa alcalina de camarón (Exo/Sap-It, USB) para eliminar el resto de cebadores y dNTPs, respectivamente. Las condiciones de secuenciación por PCR fueron 96°C 10 s, 68°C 2 min durante 25 ciclos para todos los casos, realizándose la incubación en un termociclador PTC-200 (MJ Research). A continuación, los productos de la reacción fueron purificados mediante precipitación con 2 vol de etanol 96% y 0,1 vol de acetato sódico 3 M pH 4.6 y analizados en un secuenciador ABI PRISM 377 DNA Sequencer (Perkin Elmer). El tratamiento de los datos de secuenciación se llevó a cabo empleando el programa BioEdit (Hall, 1999).

5. ANÁLISIS DE ESTRUCTURAS G-CUADRÚPLEX

Para analizar el efecto del motivo (TGGGGC)₄ sobre la síntesis de DNA se llevó a cabo un ensayo de protección de metilación con dimetilsulfato (DMS) (Han et al., 1999). Este ensayo se fundamenta en que el DMS metila los N⁷ de las guaninas los cuales están implicados en los enlaces que se establecen entre las guaninas vecinas para formar estructuras G-cuadrúplex. Por lo tanto, se encontrarán protegidos frente a la metilación con DMS. Posteriormente al tratamiento con DMS, el DNA se trata con piperidina, lo que produce cortes en los nucleótidos cuyas bases se encuentran metiladas. El análisis de estos ensayos se efectúa mediante geles de poliacrilamida. En la Figura 10 se muestra un esquema de los pasos seguidos en estos experimentos.

Los oligonucleótidos (Sigma Genosys), empleados en nuestros experimentos de detección de G-cuadrúplex fueron los siguientes:

F₁ (5'-TGAGGCTGAGGCTGGGGCTGGGGCTGGGGCTGGGGCTGAGGCTGAGGC-3')

F₂ (5'-TGAGGCTGAGGCTGGGGCTGGGGCTGGGGCTGAGGCTGAGGC-3')

F₃ (5'-TGAGGCTGAGGCTGGGGCTGAGGTGAGGCTGAGGCTGAGGCTGAGGTGAGGC-3')

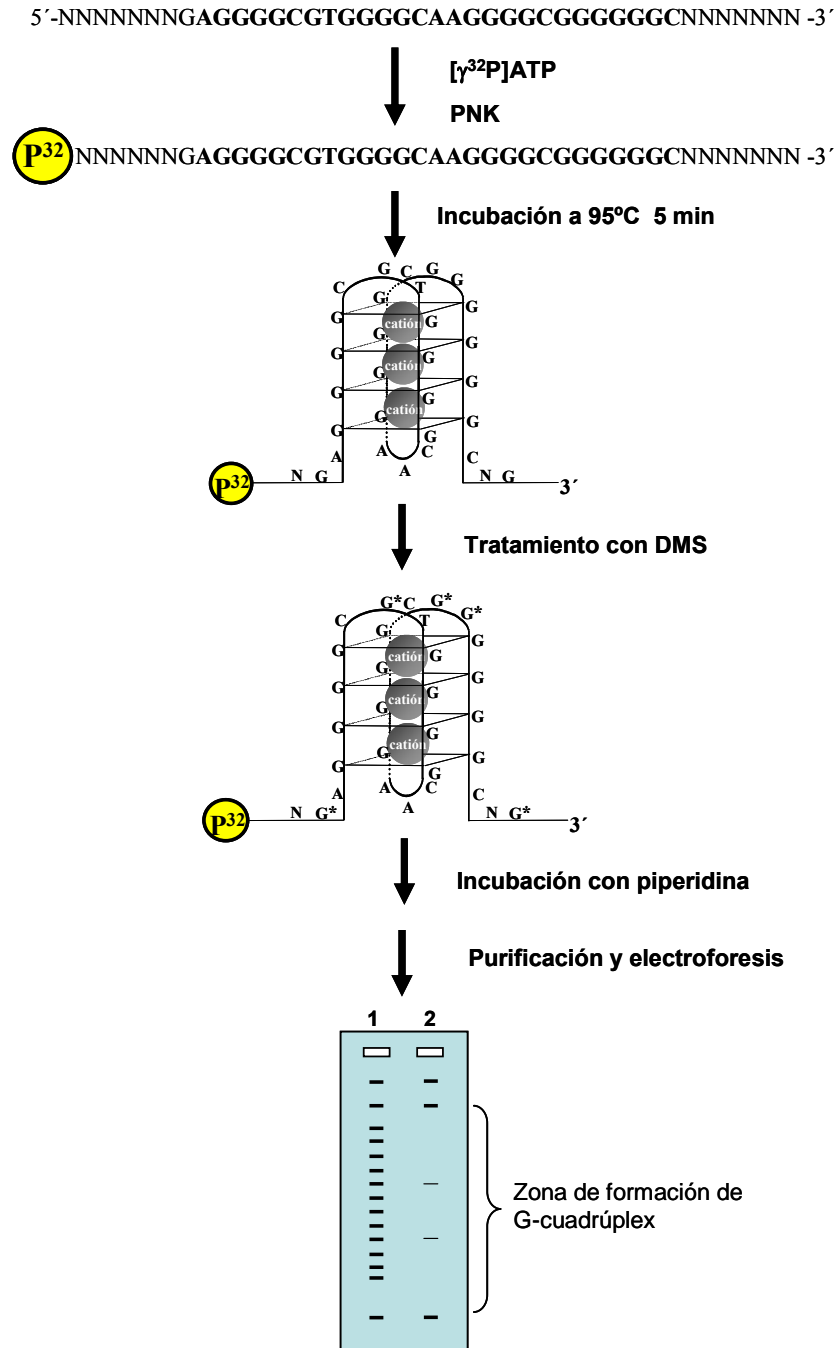


Figura 10. Representación esquemática de los experimentos de metilación con DMS para detectar estructuras G-cuadrúplex. En negrita se destaca la secuencia potencialmente formadora de estructuras G-cuadrúplex. En el esquema de una estructura G-cuadrúplex, G* representa las guaninas metiladas. En la representación del gel, el carril 1 correspondería a la muestra procedente de un ensayo sin aporte adicional de un catión, el carril 2 a la muestra de un ensayo en el que se adicionó un catión que estabiliza estructuras G-cuadrúplex.

Estos oligonucleótidos fueron marcados radiactivamente en su extremo 5' con 25 μCi de ($\gamma\text{-}^{32}\text{P}$) ATP (6000 Ci/mmol) (Amersham Biosciences) y T4 polinucleótido kinasa (PNK) (Promega) y se purificaron con el Nucleotide Removal Kit (QIAGEN). Las actividades específicas de los DNAs marcados fueron de aproximadamente 5×10^8 cpm/ μg . Con el fin de favorecer la formación de G-cuadrúplex, se incubaron 280 fmol de cada oligonucleótido marcado en 200 μl del tampón TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) conteniendo 50 o 100 mM KCl durante 5 min a 95 °C. También se realizaron incubaciones en el mismo tampón TE pero conteniendo 100 mM de NaCl, LiCl o NH_4Cl en lugar de KCl, así como también sin ningún aporte salino adicional. A continuación, los DNAs se dejaron enfriar hasta alcanzar la temperatura ambiente para luego realizar una incubación en un baño a 37 °C durante 60 min previa al tratamiento con DMS. Los oligonucleótidos fueron metilados mediante la adición de 1 μl de una dilución 1:4 de DMS en etanol e incubado durante 15 min. Las reacciones se pararon añadiéndoles 50 μl del tampón de parada (1 M β -mercaptoetanol, 1,5 M acetato sódico pH 7.0), se precipitaron con 2 vol de etanol, usando 1 μg de DNA de timo de ternera como "carrier" para favorecer la precipitación y fueron resuspendidos en 10 μl del tampón TE. A continuación, se llevó a cabo el tratamiento con piperidina. Para ello, los productos resultantes de la metilación con DMS se enfriaron incubándolos en hielo durante 5 min; seguidamente, se les añadió 150 μl de una solución de piperidina 1 M y se incubaron a 90°C durante 30 min. Posteriormente, las reacciones se depositaron en hielo durante 5 min y se precipitaron con 1 ml de n-butanol, para después añadirles 150 μl de 1% SDS y 1 ml de n-butanol, volviendo a precipitar el DNA. Finalmente, se realizaron dos precipitaciones más con 0,5 ml de n-butanol. Para el análisis de los productos de este ensayo, éstos se resuspendieron en 10 μl de tampón de carga y se separaron en un gel desnaturalizante de poliacrilamida al 18%. Una vez finalizada la electroforesis, el gel fue sometido a autorradiografía empleando películas Hyperfilm MP (Amersham Biosciences).

6. ENSAYO DE INTERFERENCIA DE LA SÍNTESIS DE DNA

El análisis del efecto que produce la presencia del motivo $(\text{TGGGGC})_4$ sobre la síntesis de DNA, se llevó a cabo la secuenciación de los plásmidos que contenían los alelos 73.1 y 80.1 de Msh43 en presencia o ausencia de KCl en la reacción. Las reacciones se realizaron en un volumen de 10 μl conteniendo 0,16 pmol de DNA

plasmídico, 1,6 pmol del cebador P02.2 y Taq DNA polimerasa. En los experimentos que contenían KCl, esta sal se añadía a la reacción de manera que la concentración final fuese de 50 mM y 62,5 mM. La secuenciación se realizó durante 24 ciclos de amplificación de 96 °C 10 s y 54 °C 2 min. A continuación, los productos de la reacción fueron purificados y analizados tal como se describió en el apartado de secuenciación.

7. ENSAYO DE HETERODÚPLEX

Para la detección de heterodúplex se mezclaron alícuotas de 5 µl del producto de amplificación de la muestra de DNA genómico con 5 µl del producto de amplificación de cada uno de los alelos de MsH43 previamente caracterizados. A continuación se incubaron cada una de las mezclas a 95 °C durante 3 min y se dejaron enfriar hasta alcanzar la temperatura ambiente. El análisis electroforético se llevó a cabo en geles de agarosa al 2% en 1X TAE (40 mM Tris-acetato, 1 mM EDTA, pH 8.0) y de poliacrilamida (29:1) al 5% en 1X TBE (90 mM Tris-borato, 1 mM EDTA, pH 8.3). Los geles se tiñeron con bromuro de etidio (0,5 µg/ml) y se visualizaron empleando un transiluminador GelDoc2000 (Biorad).

8. ENSAYOS DE RECOMBINACIÓN HOMÓLOGA INTRAMOLECULAR *IN VITRO*

8.1 Construcción de los sustratos plasmídicos

Los plásmidos utilizados en los experimentos de recombinación homóloga intramolecular *in vitro* se construyeron empleando como plásmido base el vector pBR322. En la figura 11 se indican los procesos seguidos para la preparación de los plásmidos p73.1 y p80.1 que contienen dos copias idénticas de los alelos de MsH43, 73.1 y 80.1 respectivamente, ambas en la misma orientación.

El fragmento de DNA que contiene el alelo 80.1 del minisatélite MsH43 se obtuvo mediante la digestión con *EcoRI* del plásmido pGEM-T Easy (Promega) en el que estaba clonado dicho alelo tal y como se explicó anteriormente. En el caso del alelo MsH43 73.1, el plásmido pGEM-T Easy que lo contenía se digirió con las enzimas *SacII* y *SalI*. A continuación, los insertos liberados se purificaron a partir de un gel de agarosa utilizando el QIAEX II Gel Extraction Kit (Qiagen), siguiendo las instrucciones del fabricante. Para la clonación de los alelos en el vector pBR322, este plásmido fue

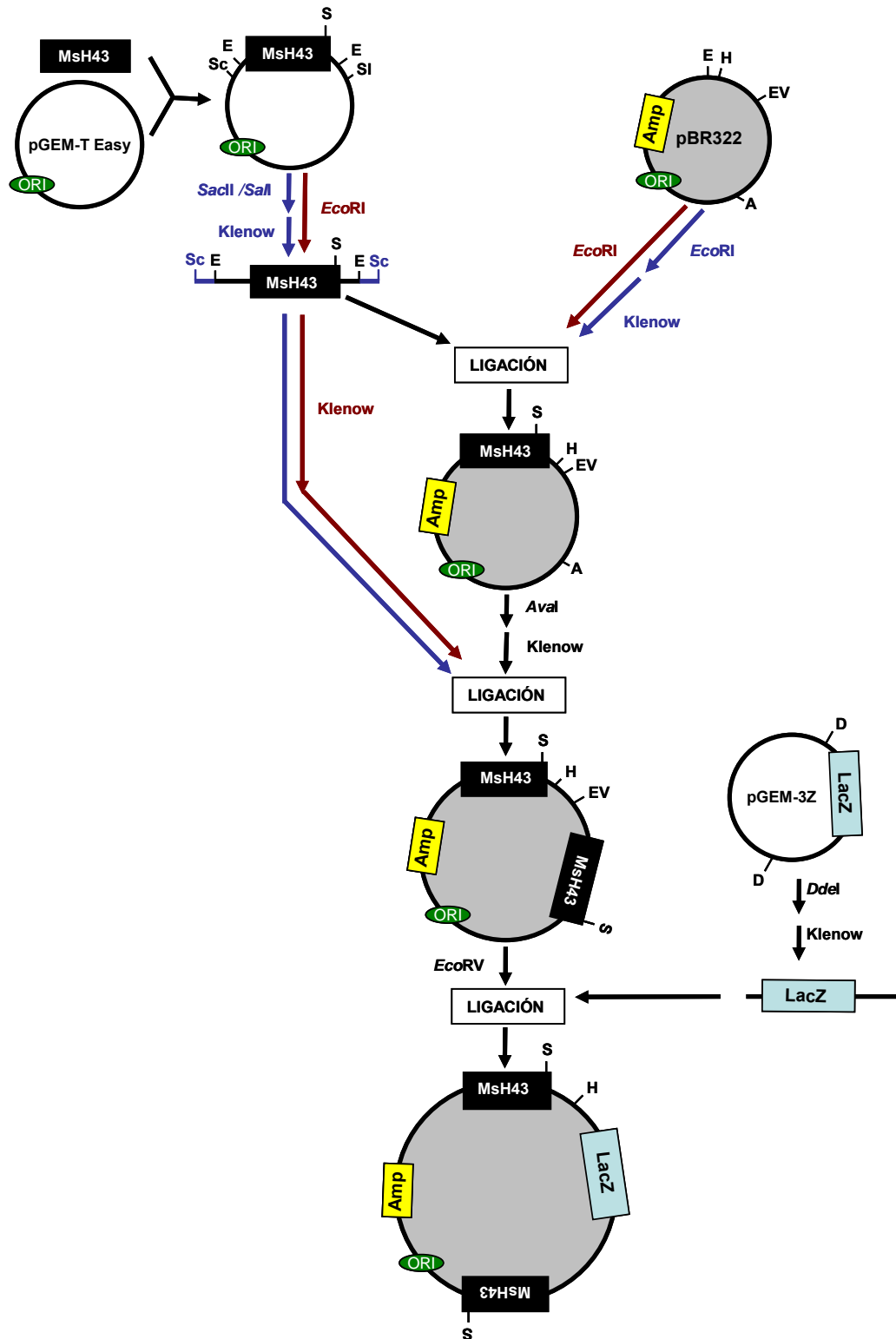


Figura 11. Representación esquemática del procedimiento seguido para la construcción de los plásmidos p73.1 y p80.1. La caja **Msh43** representa el alelo 73.1 ó 80.1 de Msh43, la caja **Amp** el gen de resistencia a la ampicilina, **ORI** el origen de replicación y la caja **LacZ** el fragmento amino terminal del gen de la β-galactosidasa, que permite el proceso de la α-complementación. **E**, *EcoRI*; **S**, *SacI*; **Sc**, *SacII*; **Sl**, *Sall*; **H**, *HindIII*; **EV**, *EcoRV*; **A**, *AvaI*; **D**, *DdeI*. En azul y rojo se indican pasos seguidos exclusivamente para la construcción de p73.1 y p80.1, respectivamente. El negro muestra procedimientos comunes en la construcción de ambos plásmidos.

previamente digerido con *EcoRI*. El alelo MsH 80.1 se clonó en pBR322 directamente, sin embargo en el caso del MsH43 73.1, como los extremos generados por los enzimas *SacII* y *SalI* no son compatibles con los que produce el enzima *EcoRI* empleado al digerir el vector pBR322, los extremos del inserto y del vector fueron rellenados con el enzima Klenow. Para ello se trataron 5 µg de DNA de ambos fragmentos con 5 U de Klenow en 50 µl de volumen final de reacción (50 mM Tris-HCl pH 7.2, 10 mM MgSO₄, 0,1 mM DTT, 20 µg/ml BSA) conteniendo 40 µM de cada dNTP. La mezcla se incubó durante 10 min a temperatura ambiente, deteniéndose la reacción por incubación durante 10 min a 75°C. El DNA fue extraído y precipitado. Para evitar la religación del plásmido, se trataron con 5 U de fosfatasa alcalina (Roche Molecular Biochemicals) 2µg de pBR322 digerido con *EcoRI* ó 2 µg de pBR322 digerido con *EcoRI* y posteriormente tratado con Klenow. El tratamiento con la fosfatasa se realizó en un volumen de reacción de 50 µl conteniendo 50 mM Tris-HCl pH 8.5 y 0,1 mM EDTA. A continuación, el DNA plasmídico fue extraído, precipitado y resuspendido en 50 µl de TE. Las reacciones de ligación entre vector e inserto se realizaron en un volumen final de 10 µl conteniendo 30 mM Tris-HCl pH 7.8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP y 5% polietilenglicol 8000, con una relación molar vector/inserto de 1/3, empleando 1 U de T4 DNA Ligasa (Promega). Las reacciones se incubaron a 4°C toda la noche y fueron inactivadas por calentamiento a 70°C durante 10 min. Posteriormente, se transformaron 100 µl de bacterias competentes *E. coli* DH5α con 2 µl de la reacción de ligación y se realizaron minipreps de las colonias obtenidas.

Para comprobar que los fragmentos de MsH43 estaban clonados en la orientación 5'→3', el DNA plasmídico obtenido de los minipreps fue digerido con los enzimas *SacI/HindIII*, cuyas dianas se encuentran en el inserto y en el plásmido respectivamente. De los plásmidos que presentaban la orientación correcta del inserto se realizó un maxiprep y se digirieron con el enzima *AvaI* para clonar una segunda copia de MsH43 80.1 y de MsH43 73.1 en el que ya contenía una copia del mismo alelo. Como los extremos generados por este enzima no permiten el clonaje directo de los insertos, todos los extremos, tanto los del vector como los de ambos insertos, fueron rellenados con el enzima Klenow tal como se indicó anteriormente. A continuación el vector fue defosforilado para evitar la religación y se realizaron las ligaciones correspondientes, siguiendo los procedimientos indicados arriba. Seguidamente, se realizaron minipreps de clones positivos, y se verificó que la orientación de los nuevos insertos fuese la misma que la del alelo de MsH43 previamente clonado en el vector,

mediante digestión con los correspondientes enzimas de restricción. Finalmente, entre las dos copias del minisatélite existentes en ambos plásmidos, se clonó el fragmento *LacZ*. Este fragmento se obtuvo mediante la digestión del plásmido pGEM3Z con el enzima DdeI. A continuación, el fragmento obtenido de esta digestión fue purificado y los extremos protuberantes se rellenaron empleando Klenow. Los plásmidos que contenían las dos copias del mismo alelo de Msh43 fueron digeridos con *EcoRV* y tratados con fosfatasa alcalina para evitar su religación antes de ser empleados en las reacciones de ligación para clonar el fragmento *LacZ*. Los clones positivos se seleccionaron por presentar α -complementación en *E. coli* DH5 α , comprobando posteriormente que la orientación de *LacZ* fuera la misma en ambos constructor.

8.2 Preparación de extractos proteicos nucleares

Los extractos proteicos nucleares se prepararon empleando el método de Dignam et al., 1983 (Ausubel et al., 1987). Como material biológico se emplearon testículos de rata Sprague-Dawley adulta (~300 g de peso, 60-65 días de edad). Los animales se sacrificaron mediante asfixia con cloroformo o CO₂ (siguiendo la normativa de la UE referente a la Experimentación y Protección Animal 86/609/CEE), tras lo cual se realizó la extracción rápida de los testículos. A partir de aquí, todos los pasos se realizaron en frío y los tampones que se emplearon fueron previamente enfriados en hielo. Los órganos extraídos se lavaron con tampón PBS (137 mM NaCl, 2,7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7,4), eliminándose los restos de sangre, tejido graso y conjuntivo. El tejido se homogenizó en PBS con un homogenizador manual de vidrio esmerilado y a continuación las células se recogieron mediante una centrifugación a 1850 g, 15 min a 4°C, midiéndose el volumen de las mismas. Este valor constituye el llamado *pcv* (*packed cell volume*). Las células se resuspendieron en un volumen de PBS igual a 5X *pcv* y se centrifugaron en las mismas condiciones que en el paso anterior. El exceso de PBS se lavó rápidamente con 5X *pcv* de tampón hipotónico (10 mM HEPES pH 7,9, 10 mM KCl, 1,5 mM MgCl₂, 0,2 mM PMSF, 0,5 mM DTT), recuperando las células por centrifugación a 1850 g, 5 min, 4°C. A continuación, las células se resuspendieron en 3X *pcv* de tampón hipotónico y se dejaron hinchar durante 10 min en hielo. Inmediatamente, las células se lisaron con el homogenizador manual para liberar los núcleos, que se recogieron por centrifugación a 3300 g, 15 min a 4°C. El volumen de los núcleos se midió para obtener el valor del *pnv* (*packed nuclei volume*) y se resuspendieron en 0,5X *pnv* de tampón de baja salinidad (20 mM HEPES pH 7,9, 20

mM KCl, 1,5 mM MgCl₂, 0,2 mM EDTA, 0,2 mM PMSF, 0,5 mM DTT, 25% glicerol). Para realizar la extracción salina de las proteínas nucleares, se añadió por goteo a la suspensión de núcleos un volumen de 0,5X *p_{nv}* de tampón de alta salinidad (20 mM HEPES pH 7,9, 1,2 M KCl, 1,5 mM MgCl₂, 0,2 mM EDTA, 0,2 mM PMSF, 0,5 mM DTT, 25% glicerol) manteniendo la mezcla en agitación durante 30 min en hielo. Seguidamente, los núcleos se centrifugaron a 25000 g, 30 min a 4°C y el sobrenadante se dializó contra 50 vol de tampón de diálisis (20 mM HEPES pH 7,9, 0,1 M KCl, 0,2 mM EDTA, 0,2 mM PMSF, 0,5 mM DTT, 20% glicerol) durante 3 h. Los restos de ácido nucleico y proteínas que precipitan durante la diálisis se eliminaron mediante una centrifugación a 25000 g, 20 min a 4°C. El sobrenadante, que constituye el extracto proteico nuclear, se alicuoteó y congeló en nitrógeno líquido, almacenándose a -70°C. La actividad del extracto bajo estas condiciones de almacenamiento permanece aparentemente inalterada durante al menos 6 meses. La concentración proteica de los extractos nucleares fue determinada mediante el método de Bradford (Bio-Rad), siendo de 4,50-5,75 µg/µl.

8.3 Ensayo de recombinación homóloga intramolecular

En los ensayos de recombinación, se siguieron las condiciones descritas por Boán et al. (1989), con ligeras modificaciones. Las reacciones de recombinación fueron efectuadas en un volumen de 100 µl que contenía 20 mM Tris-HCl pH 7,5, 10 mM MgSO₄, 1 mM ATP, 0,1 mM de cada dNTP, 10 µg de extracto proteico nuclear de testículo de rata y 1 µg de DNA plasmídico (p73.1 o p80.1). En los experimentos que contenían KCl, NaCl o NH₄Cl, la sal era añadida a la reacción de manera que la concentración final fuera de 10 mM y 20 mM. Todas las reacciones de recombinación se incubaron durante 30 min a 37 °C. A continuación, el DNA fue sometido a una extracción fenólica, con el fenol saturado con TE, seguida de otra extracción con una mezcla de fenol con cloroformo y alcohol isoamílico (25:24:1) y finalmente una extracción con una mezcla de cloroformo y alcohol isoamílico (24:1). El DNA fue precipitado con 0,1 vol de acetato sódico 3 M (pH 5,2) y 2 vol de etanol absoluto, siendo después centrifugado a 12000 g durante 15 min a 4 °C. El DNA fue resuspendido en 50 µl de Tris-HCl 10 mM pH 8,0 y conservado a -20 °C para su posterior análisis. Este producto purificado de las reacciones de recombinación se empleó para transformar 100 µl de bacterias competentes *E. coli* DH5α, que permiten la α-complementación,

utilizando 1 μ l de una dilución 1:1000 ó 1:10 dependiendo de si el experimento contenía o no sales.

En cada experimento de recombinación recuperamos dos clases de plásmidos: plásmidos originales que no habían sufrido recombinación originando colonias bacterianas de color azul (amp^r , LacZ^+), y plásmidos recombinantes que dan lugar a colonias blancas (amp^r , LacZ^-). Las colonias blancas obtenidas en la transformación se cultivaron para realizar minipreps de su DNA plasmídico. Se realizó una digestión de 300 ng del plásmido obtenido con *EcoRI* para confirmar que los plásmidos eran recombinantes, eliminando aquellas colonias que siendo blancas no correspondían con plásmidos recombinantes. Posteriormente, los plásmidos confirmados como recombinantes, eran amplificados por PCR partiendo de 1 ng de DNA plasmídico como molde siguiendo las condiciones indicadas en el apartado 3. El análisis de los productos de la amplificación se realizó mediante electroforesis en gel de agarosa al 2% que nos permitió distinguir los recombinantes iguales en tamaño al fragmento original, de aquellos que eran desiguales. Para confirmar los recombinantes iguales se llevó a cabo un ensayo de heterodúplex con estos productos, siguiendo las indicaciones mostradas anteriormente. A partir de los productos de PCR de los recombinantes elegidos se llevó a cabo la secuenciación.

RESULTADOS Y DISCUSIÓN

Análisis del polimorfismo de MsH43

La secuenciación del minisatélite MsH43, aislado a partir de una genoteca de hígado fetal humano, mostró que estaba constituido por 79 unidades repetitivas de penta y hexanucleótidos. Atendiendo al número de repeticiones que presenta y para diferenciarlos de otras posibles formas alélicas que pudiesen existir en la población humana, a este alelo se le denominó MsH43 79.1 (Figura 12). La posterior búsqueda de

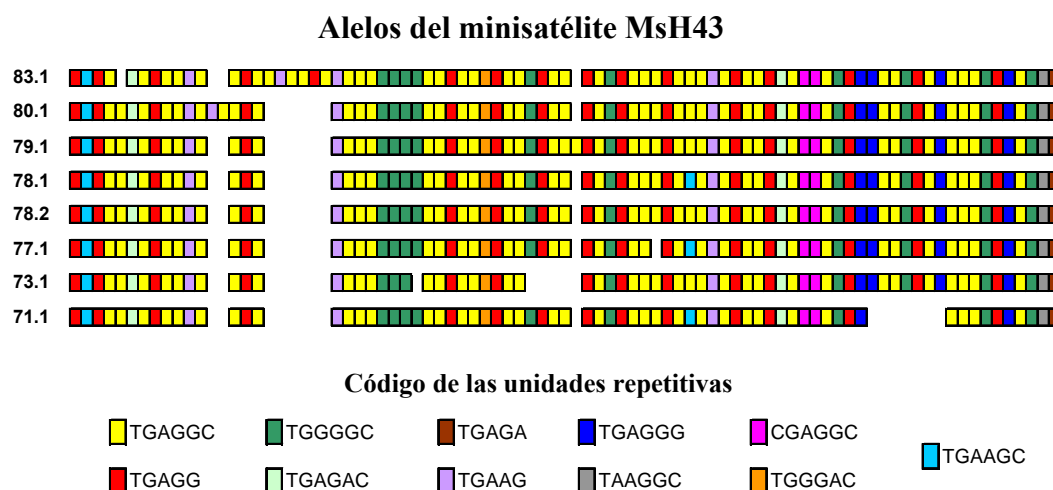


Figura 12. Representación esquemática de los diferentes alelos de MsH43 que hemos aislado en nuestro laboratorio. La denominación de cada alelo figura al margen, indicando el número de unidades repetitivas que presenta cada uno de ellos. Cada unidad repetitiva está representada por un cuadro coloreado, cuyo código también se muestra. Los espacios en blanco se han colocado para facilitar la visualización de las homología interalélicas.

MsH43 en la base de datos del genoma humano mostró que esta secuencia se encuentra localizada en el cromosoma 2, dentro de una región intergénica subtelomérica (2q37.1). La secuencia del minisatélite que figura en dicha base de datos está formada por 78 unidades repetitivas, de ahí que lo nombremos como MsH43 78.1. El análisis de la estructura de este nuevo alelo reveló que se diferencia de MsH43 79.1 por la delección de la unidad repetitiva 37 y por una transición A→G en el motivo 47.

La búsqueda e identificación de los posibles alelos se realizó mediante PCR de muestras de DNA genómico de 132 individuos caucásicos sanos no emparentados, con los cebadores P02.1 y P02.2 (Figura 13). Estos productos de PCR fueron analizados electroforéticamente mediante geles de agarosa al 2%. Aunque este análisis en geles de agarosa puso de manifiesto la existencia de polimorfismo para el locus MsH43, la asignación genotípica resultaba difícil puesto que las diferencias de tamaño entre los

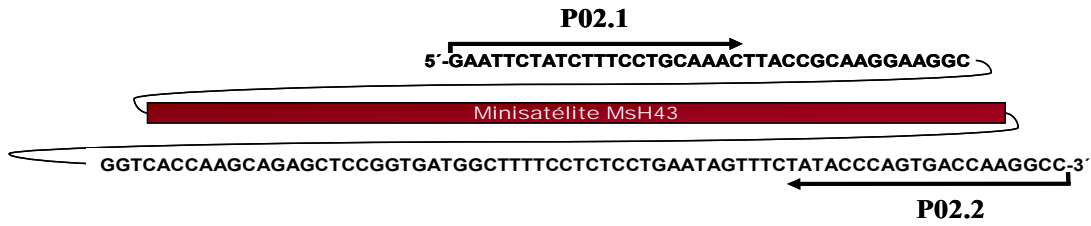


Figura 13. Representación esquemática de la posición de los cebadores P02.1 y P02.2, indicada con flechas, con respecto al minisatélite Msh43.

diversos alelos son muy pequeñas (Figura 14A). Para intentar mejorar la separación de los alelos analizamos los productos de PCR de las mismas muestras en geles de poliacrilamida al 5%. Como se muestra en la Figura 14B la poliacrilamida supuso una mejora en la resolución de las bandas que permitió detectar estructuras heterodúplex generadas durante la PCR. Dado que los heterodúplex son estructuras que se generan al hibridar cadenas distintas, esto hizo que nos planteásemos utilizar la formación de heterodúplex para la identificación de los distintos genotipos. Nuestro ensayo de heterodúplex se basa en que durante la renaturalización que tiene lugar en la PCR, las

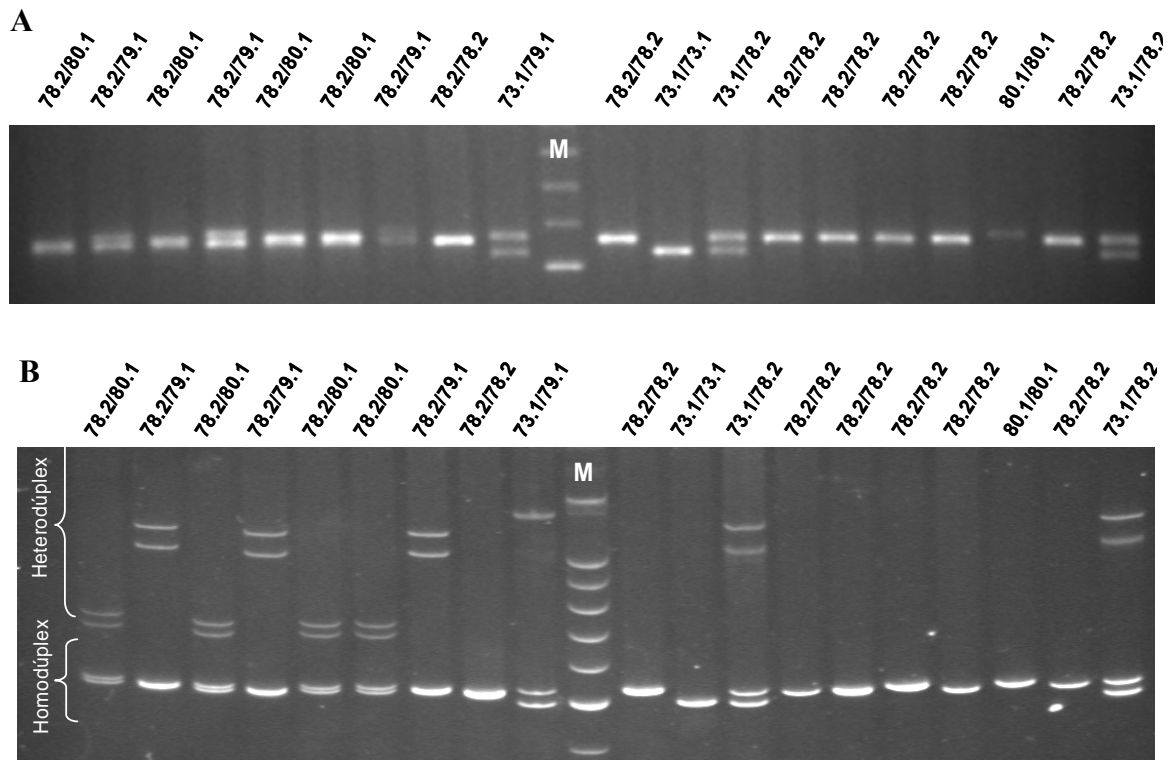


Figura 14. (A) Electroforesis en gel de agarosa al 2% de los productos de amplificación de una muestra poblacional. (B) Electroforesis en gel de poliacrilamida al 5% de los mismos productos de amplificación del gel del apartado A. Los números sobre cada carril indican los genotipos correspondientes. M, marcador 100 bp ladder de Promega.

cadenas de los dos alelos de un individuo heterocigoto, además de hibridar cada una con su complementaria dando lugar a los homodúplex correspondientes, pueden hibridar de forma cruzada produciendo moléculas heterodúplex. Estas moléculas se caracterizan porque en las zonas en las que las cadenas no son complementarias se formarán bucles de cadena sencilla, provocando un retardo en la movilidad electroforética que puede ser detectado en geles de poliacrilamida. Para corroborar si el análisis de heterodúplex podría ser un método adecuado para la identificación de alelos del minisatélite MsH43, se realizaron en primer lugar dos pruebas con alelos que ya habíamos identificado y clonado. Estos alelos eran el MsH43 79.1, que como se ha dicho había sido aislado de una genoteca de hígado fetal humano, y los alelos MsH43 73.1 y MsH43 80.1, que fueron identificados durante el análisis poblacional con los cebadores P02.1/P02.2. La primera prueba consistió en amplificar estos tres alelos con P02.1/P02.2 y analizar los productos obtenidos en geles del 5% de poliacrilamida. Este experimento permitió comprobar que la amplificación de un único alelo generaba una sola banda y que no aparecían bandas artefactuales que pudiesen llevar a una identificación alélica errónea (datos no mostrados). En la segunda prueba se realizó un análisis de heterodúplex con cada una de las tres posibles combinaciones bialélicas (73.1- 79.1, 73.1- 80.1 y 79.1- 80.1). Con este experimento se verificó que el patrón electroforético obtenido era específico de cada combinación y fácilmente reconocible (datos no mostrados).

Una vez demostrada la validez del análisis de heterodúplex para el reconocimiento de alelos de MsH43, comenzamos su aplicación al estudio del polimorfismo de este locus en la población que habíamos muestreado. Para ello, los patrones electroforéticos de los productos de amplificación de los diferentes DNA genómicos fueron comparados con los que producían los alelos previamente caracterizados. La aparición de un patrón diferente indicaba la presencia de al menos un nuevo alelo de MsH43 (Figura 15). La posterior extracción del gel de las bandas correspondientes a los homodúplex, nos permitió purificar, clonar y secuenciar los fragmentos de DNA correspondientes a las nuevas variantes alélicas de MsH43 que estábamos detectando.

Con respecto a la caracterización de los individuos homocigotos, los cuales no forman heterodúplex durante las amplificaciones de su DNA genómico, se indujo la formación de heterodúplex mezclando su producto de amplificación con el de cada uno de los alelos ya caracterizados. De esta forma, identificábamos de qué alelo se trata porque la única mezcla en la que no se detecta la presencia de heterodúplex es aquella

en la que el alelo es incubado consigo mismo. Además, los patrones electroforéticos generados por ese alelo con todos los demás coinciden con los esperados para esas

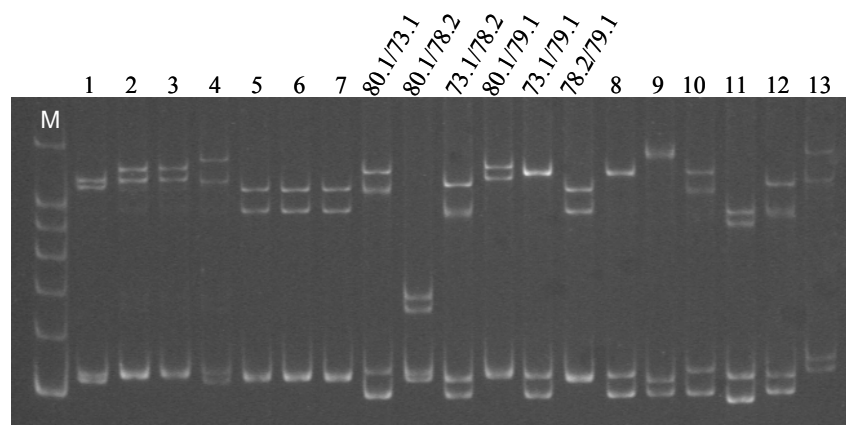


Figura 15. Electroforesis en gel de poliacrilamida al 5% en las que se muestra la formación de heterodúplex de los productos de amplificación de 13 individuos (1-13). Mediante el ensayo de heterodúplex. Además de estas muestras, también están presentes las mezclas bialélicas de los alelos ya identificados (80.1/73.1, 80.1/78.2, 73.1/78.2, 80.1/79.1, 73.1/79.1, 78.1/79.1), cuyos patrones electroforéticos se utilizan para identificar a los genotipos desconocidos mediante comparación. Así, en los carriles 1, 4, 8, 9, 11 y 13 se encuentran muestras cuyos patrones electroforéticos son diferentes a los que proporcionan los genotipos conocidos, indicando que en cada una hay al menos un alelo que todavía no ha sido caracterizado. M, marcador *100 bp ladder* de Promega.

combinaciones. Como se observa en la Figura 16 la única combinación en la que no se forman heterodúplex es cuando el alelo se mezcla consigo mismo. Además, los patrones obtenidos en las otras combinaciones coinciden con los esperados.

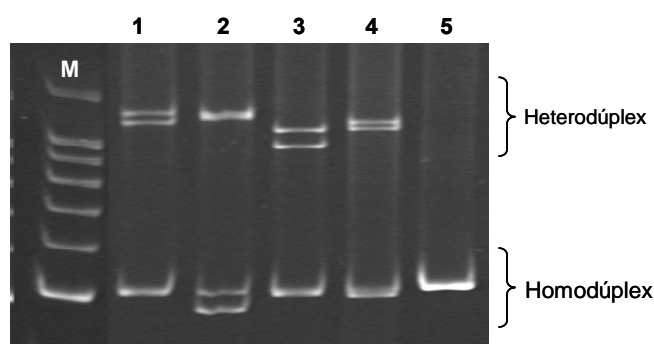


Figura 16. Electroforesis en gel del 5% de poliacrilamida en el que se muestra el análisis de heterodúplex de un producto de amplificación procedente de un individuo homocigoto para el alelo Msh43 79.1. Los carriles 1-5 representan las mezclas del producto de amplificación de ese individuo homocigoto con los productos de amplificación de los alelos de Msh43: 80.1, 73.1, 78.2, 77.1 y 79.1, respectivamente. En este caso se confirma que el individuo poseía el alelo Msh43 79.1. M, marcador *100 bp ladder* de Promega.

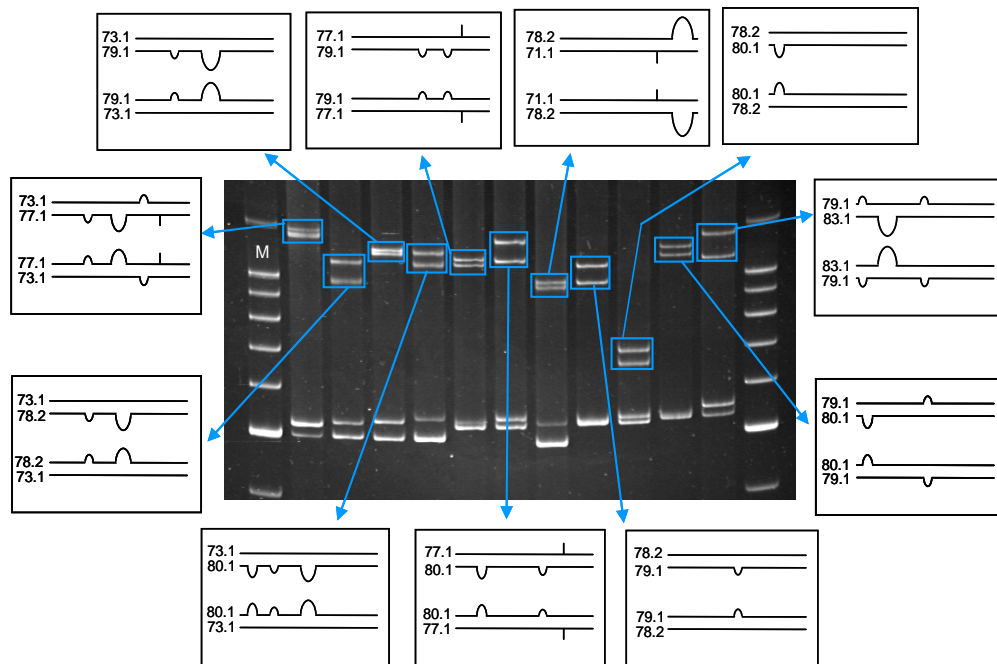
Cuando la muestra ensayada forma heterodúplex con todos los alelos conocidos, generando además patrones diferentes a los ya detectados, indica que nos encontramos

ante un nuevo alelo. En estos casos, el producto de amplificación fue clonado y secuenciado. Este estudio del polimorfismo del locus MsH43 mediante análisis de heterodúplex nos permitió identificar cuatro nuevos alelos que denominamos, siguiendo el criterio indicado anteriormente, MsH43 71.1, MsH43 77.1, MsH43 78.2 y MsH43 83.1. Estos alelos, junto con los previamente identificados, se muestran en la Figura 12.

Los resultados que obtuvimos al realizar los análisis de heterodúplex, y en concreto la movilidad de las bandas correspondientes a los fragmentos que los contenían, planteaba la siguiente pregunta: ¿Cómo son las estructuras de los bucles de cadena sencilla de los heterodúplex que alteran su movilidad electroforética? Después de secuenciar los alelos aislados y de comparar sus secuencias detectamos zonas en las que no hay complementariedad y por lo tanto donde se forman bucles de cadena sencilla. Además, concluimos que el retardo migratorio de las bandas de heterodúplex con respecto a las homodúplex, dependía de la posición del bucle ya que los heterodúplex que presentaban un menor retardo en la migración carecían de un bucle central (Figura 17). En la Figura 17A se muestran los patrones electroforéticos generados por todas las combinaciones alélicas que hemos detectado en nuestro estudio y en la Figura 17B se muestran los patrones de combinaciones alélicas que no han sido encontradas. El efecto de la posición del bucle sobre la migración electroforética ha sido descrito por otros investigadores que, al igual que en nuestro estudio, observaron que los heterodúplex que migran menos retrasados respecto al homodúplex son aquellos que no presentan un bucle central (Bhattacharyya & Lilley, 1989).

En otros casos, la aplicación del análisis de heterodúplex para el estudio del polimorfismo de ciertos loci, ha dado lugar a patrones electroforéticos complejos (Spink et al, 2004; Rodríguez et al, 2001), que dificultan la adscripción genotípica. Sin embargo, este método ha demostrado su efectividad al ser empleado en estudios taxonómicos (Kreis & Whistler, 1997), así como en el análisis del polimorfismo de secuencias microsatélite (Haddad et al, 1997). Los resultados obtenidos en nuestro trabajo ponen de manifiesto que este tipo de análisis constituye un método eficaz, simple y rápido para detectar y caracterizar variabilidad en MsH43 y que podría serlo también para analizar el polimorfismo en otros minisatélites.

A



B

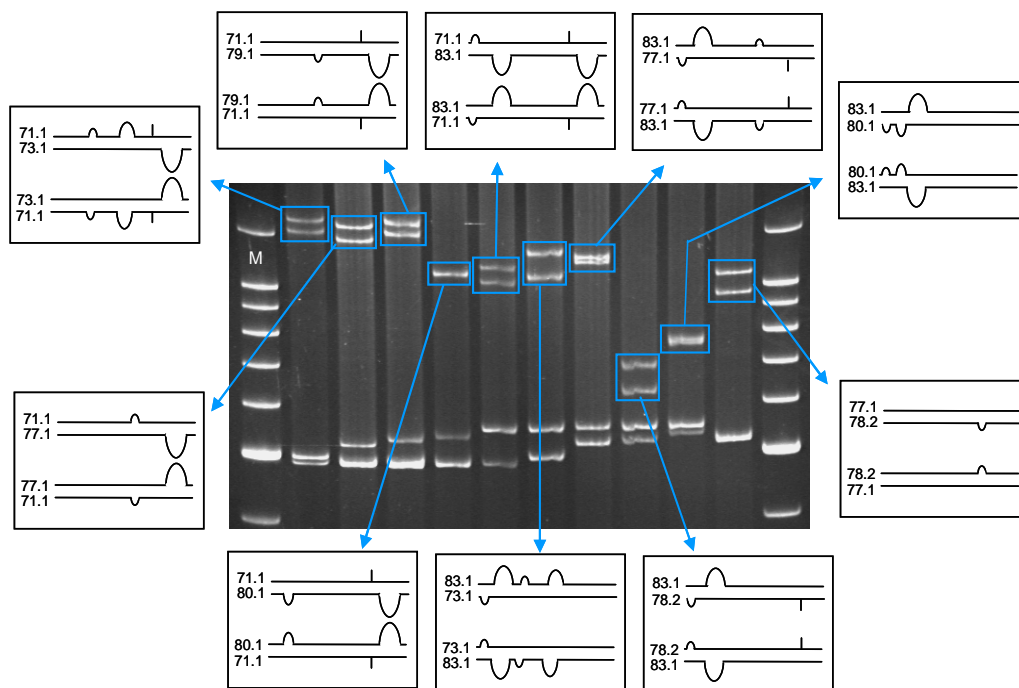


Figura 17. Estructuras heterodúplex generadas por los alelos de Msh43. Electroforesis en geles de poliacrilamida al 5% mostrando los patrones electroforéticos obtenidos en los experimentos de análisis de heterodúplex (panel A), y los resultados de las combinaciones alélicas no encontradas en la población (panel B). Asociado a cada banda heterodúplex se muestra su representación esquemática, indicando la posición de los bucles generados por la no complementariedad. Las líneas verticales señalan la presencia de mutaciones puntuales. Los números empleados en estas representaciones indican los diferentes alelos de Msh43 utilizados. M, marcador 100 bp ladder de Promega.

Amplificación preferencial y G-cuadrúplex

Al inicio de la puesta a punto de las amplificaciones por PCR para el análisis del polimorfismo del minisatélite MsH43, nos encontramos con un resultado sorprendente: las amplificaciones mostraban la existencia de polimorfismo, ya que observábamos bandas de distinto tamaño, pero no se encontraban individuos heterocigotos, es decir, con dos bandas (Figura 18). La razón más probable para este resultado es que se

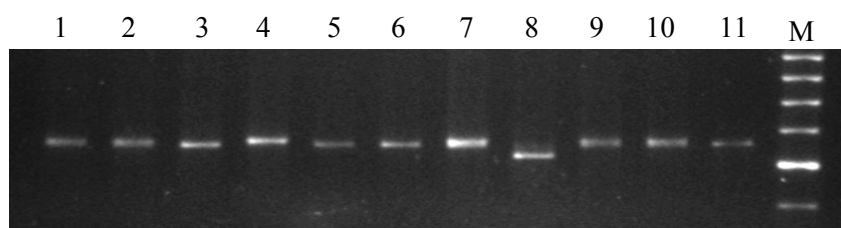
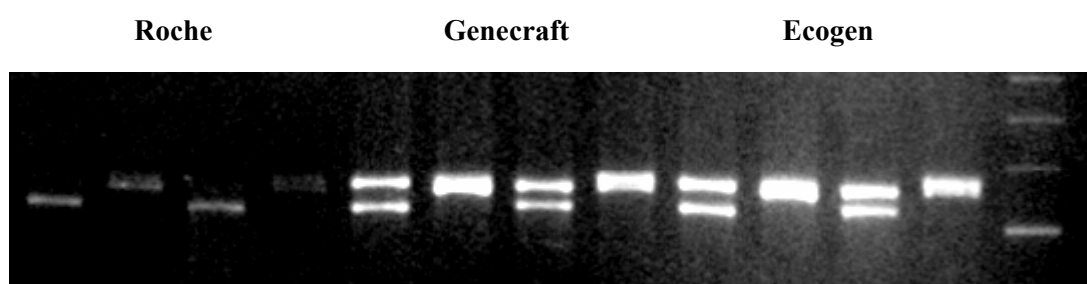


Figura 18. Electroforesis en gel de agarosa al 2% de los productos de amplificación del locus MsH43 a partir del DNA genómico de once individuos (1-11). La PCR se realizó empleando la *Taq* DNA polimerasa de Roche y su correspondiente tampón. M, marcador 100 bp ladder de Promega.

estuviese produciendo un artefacto durante la PCR, y dado que aparecía sólo una banda debía ser una amplificación preferencial. Para averiguar si la causa de este artefacto era un problema asociado a la *Taq* DNA polimerasa que se estaba empleando (*Taq* DNA polimerasa de Roche), se amplificó el genómico de cuatro individuos de la familia de referencia CEPH 1347 (01, 02, 06 y 07; padre, madre, hijo e hija, respectivamente) empleando *Taq* DNA polimerasa de otros proveedores, Genecraft y Ecogen, con sus correspondientes tampones. Los resultados de estos experimentos (Figura 19A) demostraron que cuando se empleaban las *Taq* DNA polimerasas de Genecraft y Ecogen, no sólo mejoraba la intensidad de las bandas sino que aparecían los heterocigotos (véase individuos 01 y 06). Por el contrario, la *Taq* de Roche producía únicamente la amplificación correspondiente al alelo de menor tamaño (Figura 19A). Este efecto era perfectamente reproducible en experimentos realizados con lotes diferentes de *Taq*.

Ya que las polimerasas de los demás fabricantes amplificaban perfectamente el locus Msh43, parecía que la amplificación preferencial de uno de los alelos al emplear la *Taq* de Roche estaba asociada a las características del propio enzima o a alguno de los componentes del tampón usado. Para averiguar cuál de estos dos factores era el responsable del artefacto que estábamos observando, se realizaron amplificaciones del locus Msh43 del individuo heterocigoto 01 de la familia CEPH 1347, en adelante individuo 01, en las que se empleó la *Taq* de Roche con los tampones de Genecraft e

A



B

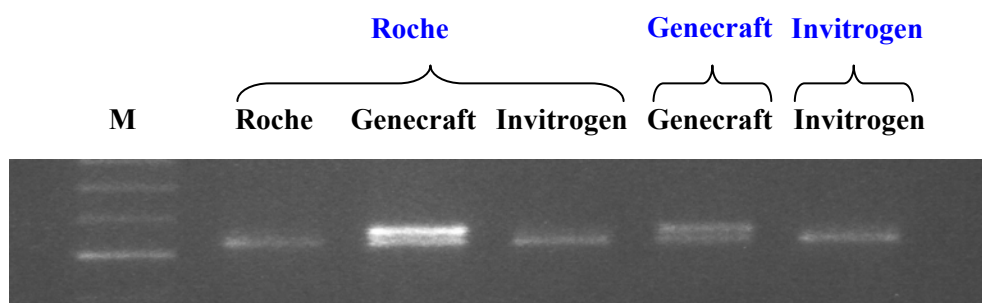


Figura 19. (A) Análisis comparativo de *Taq* polimerasas. Electroforesis en gel de agarosa al 2% de los productos de PCR obtenidos con la *Taq* DNA polimerasa de Roche, Genecraft y Ecogen. Los DNA genómicos empleados en la amplificación del locus Msh43 procedían de cuatro individuos de la familia CEPH 1347 (01, padre; 02, madre; 06, hijo y 07, hija). M, marcador *100 bp ladder* de Promega. (B) Electroforesis en gel de agarosa al 2% en la que se muestra la amplificación con la *Taq* DNA polimerasa de Roche empleando tres tampones de PCR de diferentes casas comerciales. En azul, se indica la marca del enzima y en negro el tampón utilizado. M, marcador *100 bp ladder* de Promega.

Invitrogen. El resultado de este ensayo mostró que la DNA polimerasa de Roche era capaz de amplificar perfectamente los dos alelos únicamente cuando el tampón de

reacción empleado era el de Genecraft (Figura 19B). Estos datos demostraban que la amplificación preferencial se debía a la composición del tampón utilizado y no al enzima.

Una vez establecido que el componente implicado en la generación del artefacto estaba en el tampón, para averiguar cuál era ese componente que provocaba la amplificación preferencial se prepararon tampones con la misma composición que los de Roche (tampón 1) y Genecraft (tampón 7) y una serie de tampones en los que se fueron combinando características de uno y otro tampón (ver Tabla 4 en la sección de material y métodos). La *Taq* DNA polimerasa empleada en las PCRs que se realizaron con estos doce tampones fue la de Roche y el DNA empleado en estos experimentos fue el del individuo 01, que era heterocigoto para el locus *Msh43*. Los resultados de estos experimentos se muestran en la Figura 20. Puede apreciarse como los tampones

Carril	G	R	1	2	3	4	5	6	7	8	9	10	11	12
Tris-HCl (mM)	67	10	10	67	10	10	10	10	67	10	67	67	67	67
pH	8.8	8.3	8.3	8.3	8.8	8.3	8.3	8.3	8.8	8.8	8.3	8.8	8.8	8.8
KCl (mM)	0	50	50	50	50	0	50	50	0	0	0	50	0	50
(NH ₄) ₂ SO ₄ (mM)	16	0	0	0	0	16	0	16	16	16	16	0	16	16
Tween-20 (%)	0.01	0	0	0	0	0	0.01	0	0.01	0.01	0.01	0.01	0	0.01

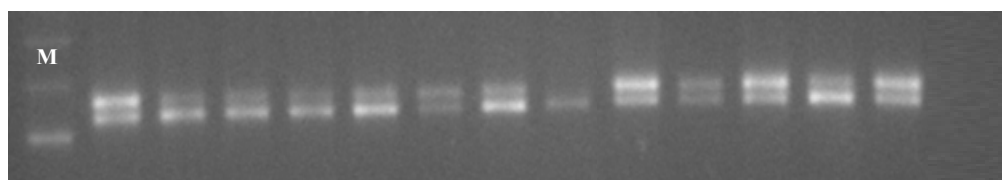


Figura 20. Efecto de la composición del tampón en la amplificación preferencial. Electroforesis en geles de agarosa al 2% de los productos de amplificación del locus *Msh43* del individuo 01 de la familia CEPH 1347 realizadas con la *Taq* de Roche empleando diferentes tampones cuya composición está indicada sobre cada carril. G y R simbolizan los tampones comerciales de Genecraft y Roche, respectivamente.

suministrados por las casas comerciales de Roche y Genecraft y los tampones 7 y 1, idénticos a los anteriores pero preparados en nuestro laboratorio, proporcionan los mismos resultados (comparar carriles G y 7 y carriles R y 1). De nuevo se pone de manifiesto que cuando la amplificación se realiza con la *Taq* de Roche y su tampón, ya sea el comercial o el preparado en nuestro laboratorio, se amplifica de forma preferencial el alelo de menor tamaño. Del mismo modo, cuando se emplea la *Taq* de Genecraft y su tampón, el comercial o el preparado en nuestro laboratorio, se obtiene la misma intensidad en la amplificación de ambos alelos (Figura 20). Los datos obtenidos

de este análisis indican que siempre que se encuentra presente KCl en el tampón, existe un sesgo hacia una amplificación preferencial del alelo menor (ver carriles R, 1, 2, 3, 5, 6 y 10), mientras que cuando la sal empleada es $(\text{NH}_4)_2\text{SO}_4$, la proporción de amplificación de ambos alelos tiende a equilibrarse (ver carriles G, 4, 7, 8, 9 y 11). El hecho de que con los tampones 6 y 12 se obtengan unas amplificaciones débiles o nulas, posiblemente sea debido a la elevada fuerza iónica de la disolución ya que en dichos tampones coexisten ambas sales. Por otro lado, se aprecia como el efecto del pH es poco relevante (comparar carriles G y R con carriles 9 y 3, respectivamente) y que la presencia de $(\text{NH}_4)_2\text{SO}_4$ hace necesaria una mayor concentración de Tris-HCl para mantener la eficacia de la *Taq* (comparar carriles 7 y 8). Por lo tanto, estos resultados sugerían que el KCl podía ser el responsable de la amplificación preferencial.

Para confirmar la hipótesis de que la presencia de KCl era la responsable de la amplificación preferencial se elaboraron ocho tampones con concentraciones crecientes de KCl (Tabla 4, tampones 13-20) y se realizaron amplificaciones con el DNA de la muestra 01. Los resultados de las reacciones de PCR (Figura 21A) mostraron que entre 25 y 37,5 mM de KCl la amplificación de ambos alelos era muy similar, mientras que por encima de 50 mM se apreciaba una clara diferencia a favor del alelo de menor tamaño. Concentraciones de KCl menores de 12,5 mM y mayores de 87,5 mM producían la inactivación de la polimerasa, como consecuencia de la variación (por exceso o por defecto) de la fuerza iónica presente en la reacción. El hecho de que al aumentar el KCl se detectase la amplificación preferencial planteaba la siguiente cuestión: ¿es el efecto del KCl específico de esta sal o una consecuencia de cambios en la fuerza iónica del tampón? Para responder a esta pregunta se realizaron PCR empleando el mismo DNA genómico pero usando tampones en los que la concentración de KCl iba disminuyendo, siendo sustituido por $(\text{NH}_4)_2\text{SO}_4$, de manera que se iba reduciendo la concentración del potasio sin variar la fuerza iónica (ver Tabla 4 en la sección de material y métodos). Como se muestra en la Figura 21B, a concentraciones elevadas de KCl se obtenía una amplificación más débil, con una clara amplificación preferencial del alelo de menor tamaño. A medida que sustituíamos el KCl por $(\text{NH}_4)_2\text{SO}_4$, la intensidad de las dos bandas se equiparaba, llegando a un óptimo de amplificación en la relación 6 mM KCl//14 mM $(\text{NH}_4)_2\text{SO}_4$. Si elevábamos la concentración de $(\text{NH}_4)_2\text{SO}_4$ en exceso, la actividad del enzima decrecía, al igual que sucedía en los experimentos variando la concentración del KCl (comparar últimos carriles en Figuras 21A y 21B).

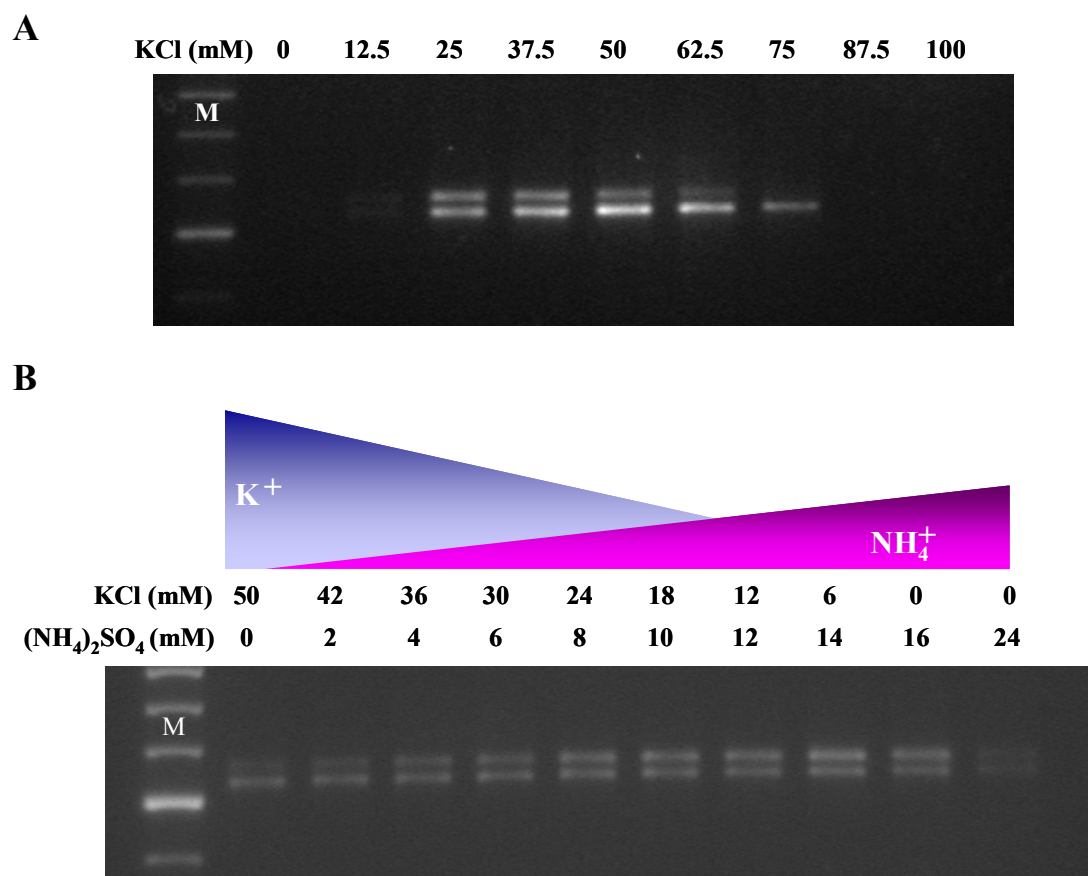


Figura 21. Efecto de la presencia de potasio en la amplificación preferencial. Electroforesis en geles de agarosa al 2% de los productos de amplificación del locus MsH43 del individuo 01 de la familia CEPH 1347 realizadas con la *Taq* de Roche empleando diferentes tampones. **(A)** Amplificaciones realizadas con los tampones 13, 14, 15, 16, 10, 17, 18, 19, y 20 (tabla 4), que se diferencian únicamente en la concentración de KCl que contienen, la cual se indica sobre cada carril. **(B)** Amplificaciones realizadas con los tampones 10, 21, 22, 23, 24, 25, 26, 27, 8 y 28 (tabla 4) que se elaboraron sustituyendo el KCl por el (NH₄)₂SO₄ en su composición. M, marcador 100 bp ladder de Promega.

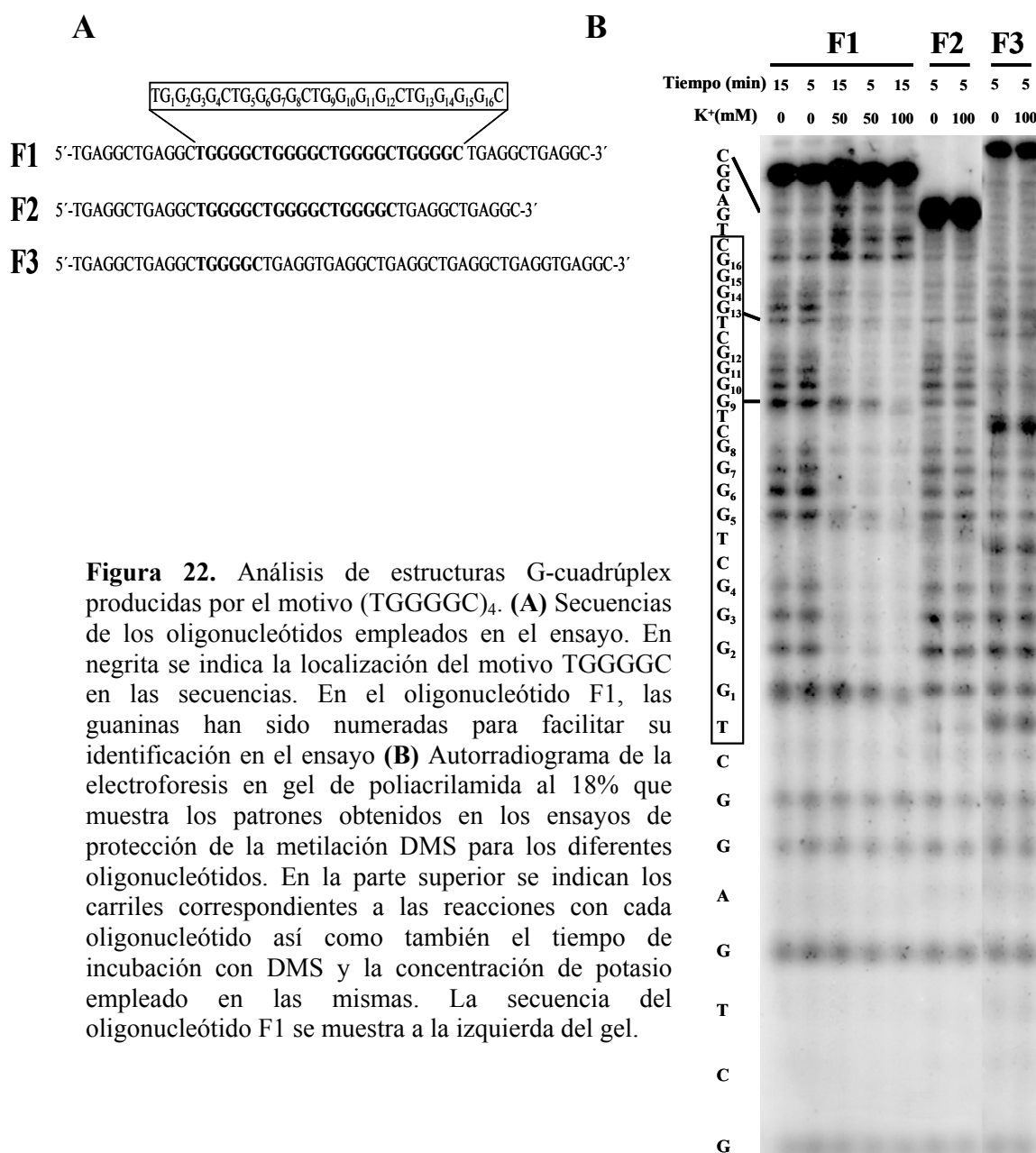
Los resultados de estos experimentos nos permiten concluir que la amplificación preferencial observada en el locus MsH43 era provocada por la presencia específica de KCl en las reacciones.

Este resultado planteaba una nueva cuestión: ¿cuál es la relación entre la secuencia de los alelos y la amplificación preferencial observada en un medio rico en potasio? Para dar respuesta a esta cuestión los dos alelos de la muestra 01 fueron clonados y secuenciados. Como se mencionó en el bloque anterior, el alelo de menor tamaño es de 525 pb y está formado por 73 unidades repetitivas, de ahí que se le denominase MsH43 73.1. Por su parte, el alelo de mayor tamaño es de 565 pb y presenta 80 repeticiones, denominándosele MsH43 80.1. La comparación de las secuencias de ambos alelos (Figura 12) muestra que su contenido en GC (66.0% para

Msh43 73.1 y 66.2% para Msh43 80.1) es similar. Las diferencias más salientables que se observan entre ambas secuencias son las siguientes: i) el alelo Msh43 80.1 contiene una duplicación de las unidades repetitivas 11 y 12, que está ausente en el Msh43 73.1; ii) el alelo Msh43 73.1 carece de un bloque de cinco unidades repetitivas localizado entre las repeticiones 35 y 39 del alelo Msh43 80.1; iii) en el alelo Msh43 80.1 hay cuatro repeticiones de la unidad TGGGGC, mientras que el Msh43 73.1 presenta sólo tres. Esta última diferencia fue la que llamó nuestra atención, ya que el motivo TGGGGC cuando está cuadruplicado $-(TGGGGC)_4-$ puede ser una secuencia potencialmente formadora de estructuras G-cuadrúplex (Simonsson, 2001). Este tipo de estructuras pueden interferir en la síntesis de DNA (Woodford et al., 1994) pudiéndose producir, por lo tanto, una amplificación preferencial. Otra prueba que apuntaba a que la presencia de $(TGGGGC)_4$ era el responsable de la amplificación preferencial era el hecho de que este tipo de estructuras son estabilizadas por cationes monovalentes, siendo el potasio el más importante (Sen & Gilbert, 1990). Además, ha sido descrito que al emplear tampones sin potasio se reduce de manera notable la generación de extensiones incompletas durante las amplificaciones y secuenciaciones por PCR (Woodford et al, 1995). Por lo tanto, si el motivo $(TGGGGC)_4$ presente en el alelo Msh43 80.1 tuviese la capacidad de formar estructuras cuadrúplex y estas ser estabilizadas por potasio, se podría explicar porqué en presencia de este ión se produce la amplificación preferencial del alelo Msh43 73.1, ya que al carecer éste de la secuencia formadora de G-cuadrúplex no generaría dichas estructuras y no habría interferencia en la síntesis de DNA durante la amplificación de este alelo.

Para verificar que el motivo $(TGGGGC)_4$ presente en el alelo Msh43 80.1 formaba estructuras G-cuadrúplex en presencia de potasio se realizó un ensayo de protección de la metilación por DMS. El fundamento de este experimento se basa en que el DMS metila los N⁷ de las guaninas, los cuales están implicados en el establecimiento de los puentes de hidrógeno que, junto con el catión, estabilizan las estructuras G-cuadrúplex. Por lo tanto, si dicha estructura se establece, el DMS no puede metilar a las guaninas implicadas en su formación. Como sustrato para estos experimentos, se diseñaron tres oligonucleótidos denominados F1, F2 y F3 de 48, 42 y 52 pb, respectivamente (Figura 22A). La secuencia de F1 contiene 8 unidades repetitivas presentes en el alelo Msh43 80.1 que contenían el motivo $(TGGGGC)_4$ en posición central. Por su parte, F2 tiene la misma secuencia que F1, excepto por la ausencia de una de las unidades TGGGGC de dicho motivo, al igual que ocurre en el

alelo MsH43 73.1. Como secuencia control en nuestros experimentos, utilizamos el oligonucleótido F3, constituido por 9 unidades repetitivas de la parte central de MsH43 80.1. En la Figura 22B se muestra el resultado de los experimentos de protección de la metilación por DMS, tanto en presencia como en ausencia de potasio, realizados con los tres oligonucleótidos marcados isotópicamente.



Este ensayo demostró que las guaninas contenidas en el motivo (TGGGGC)₄ del oligonucleótido F1 estaban protegidas frente a la metilación con DMS cuando estaba presente el potasio (véanse los carriles 50 y 100 mM KCl de F1 en la figura 21A). Sin embargo, en ausencia de este ión las guaninas eran metiladas. En las reacciones en las

que se usaron F2 y F3 no se observó ninguna protección frente a la metilación ni en ausencia o presencia potasio lo que indicaba que para formar cuadrúplex era absolutamente necesaria la presencia del motivo TGGGGC repetido cuatro veces. Estos resultados eran coherentes con la hipótesis de la formación de estructuras cuadrúplex en las que los N⁷ de las guaninas estarían protegidas. Por otra parte, el hecho de que las guaninas G1, G5, G9 y G13 fuesen parcialmente susceptibles de ser metiladas, sugería el posible modelo de G-cuadrúplex que estaría adoptando la secuencia. En este modelo, dichas guaninas no estarían involucradas en los puentes de hidrógeno que estabilizan el G-cuadrúplex sino que se encontrarían situadas en los bucles de la estructura por lo que serían más accesibles al reactivo DMS (Figura 23).

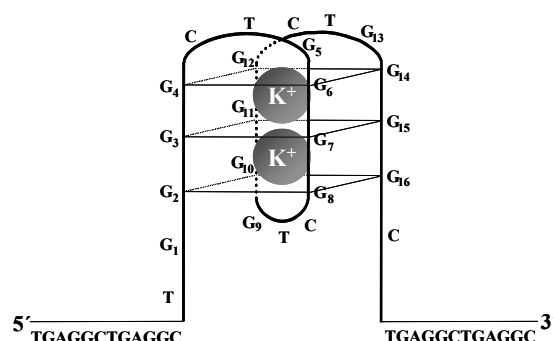


Figura 23. Representación esquemática de una posible conformación del G-cuadrúplex detectado en el locus Msh43 (modificado de Han et al, 1999).

Se ha descrito que diversos cationes presentan capacidad de estabilizar estructuras G-cuadrúplex. Entre estos se encuentran el sodio (Crnugelj et al., 2002), el amonio (Hud et al., 1998), el estroncio (Chen, 1992) o el plomo (Smirnov et al., 2000), aunque, como se ha dicho, el catión que ha mostrado ser más efectivo es el potasio (Sen & Gilbert, 1990). También existen G-cuadrúplex que pueden ser estabilizados por distintos cationes, aunque en algunos casos la conformación del cuadrúplex resultante es distinta dependiendo de cuál sea el catión estabilizador de la estructura (Weitzmann et al., 1998). Para verificar si en el caso de Msh43 el G-cuadrúplex pudiera ser estabilizado por otros cationes, se llevaron a cabo reacciones de protección de la metilación con DMS empleando sodio, litio y amonio. En estos experimentos, en los que en vez de potasio los oligonucleótidos se incubaron con estos otros iones, no se detectó protección contra la metilación (datos no mostrados), lo que indicaba que estos iones no estabilizan la estructura G-cuadrúplex que se establece en el motivo (TGGGGC)₄ del minisatélite Msh43. La especificidad que muestra el potasio para

estabilizar estas estructuras en MsH43, también ha sido observada con otras secuencias potencialmente formadoras de G-cuadrúplex (Woodford et al., 1994).

Una vez demostrado que el motivo (TGGGGC)₄ es capaz de formar G-cuadrúplex que son estabilizados específicamente por potasio, nos planteamos comprobar si estas estructuras interfieren sobre la síntesis de DNA. Para ello se realizaron reacciones de secuenciación por PCR que consistían en secuenciar los alelos MsH43 73.1 y MsH43 80.1 en ausencia de KCl y con 50 y 62,5 mM de esta sal. Los resultados de estos experimentos se muestran en la Figura 24. Como se puede apreciar,

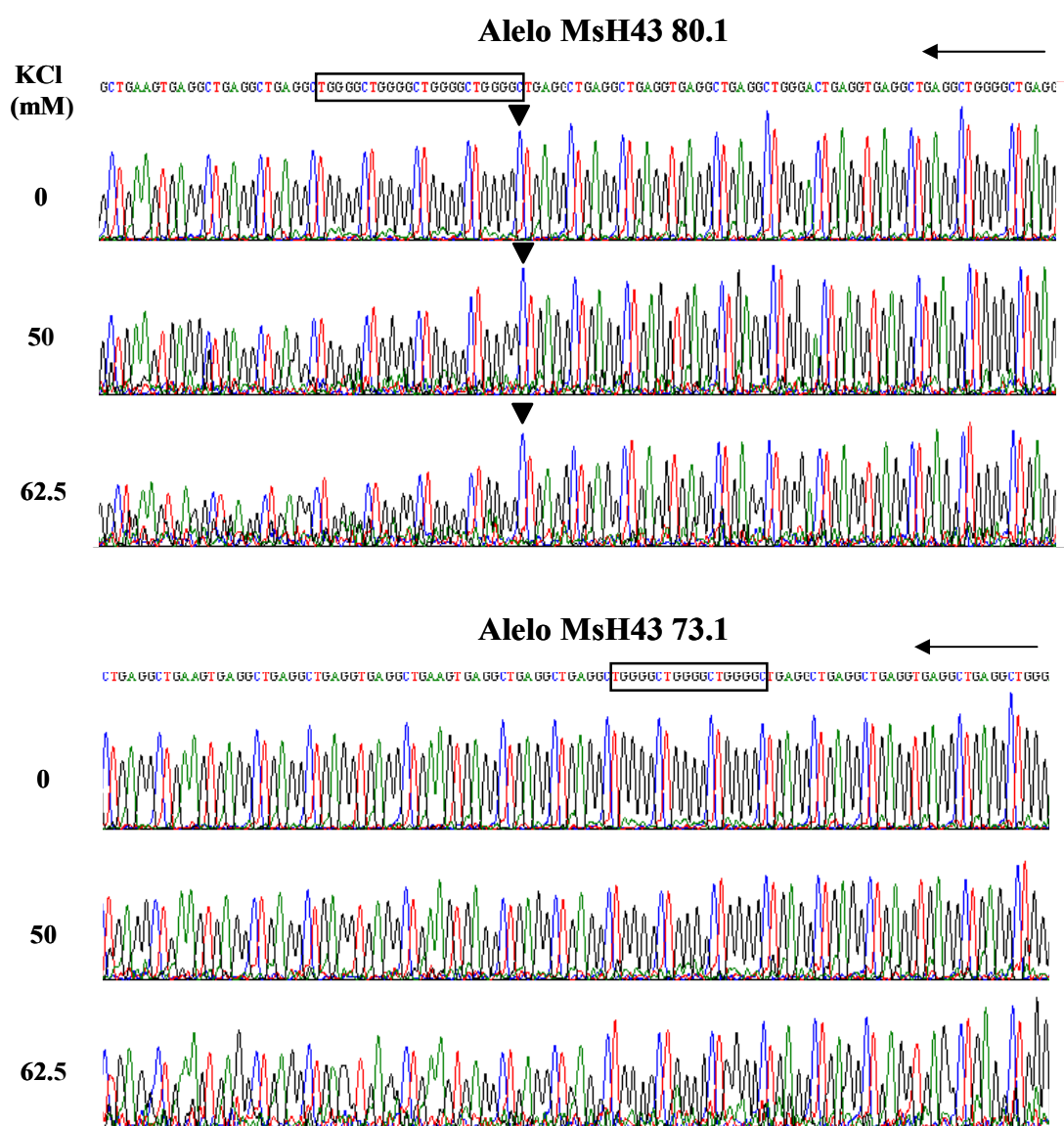


Figura 24. Análisis de interferencia en la síntesis de DNA producida por la generación de G-cuadrúplex. Se muestran los electroferogramas del análisis de interferencia de la síntesis de DNA realizados con los alelos 73.1 y 80.1 de MsH43, a las concentraciones de KCl indicadas. Las cajas indican la posición de la secuencia que contiene el motivo TGGGGC. En los electroferogramas correspondientes al alelo MsH43 80.1 se señala con una punta de flecha el inicio de la interferencia en la síntesis de DNA. Es de notar que esta posición coincide con el inicio del el motivo (TGGGGC)₄. La flecha indica la dirección de la secuencia.

los electroferogramas de dichas secuenciaciones ponen de manifiesto como la síntesis de DNA se reducía notablemente a partir del motivo (TGGGGC)₄ presente en el alelo Msh43 80.1. Este efecto se apreciaba ligeramente en ausencia de potasio pero era muy claro cuando este ión estaba presente, tanto a una concentración de 50 mM, como de manera aún más pronunciada, a 62,5 mM de KCl. Por el contrario, el alelo Msh43 73.1 no sufría ninguna reducción en la síntesis de DNA. Estos resultados demuestran que la interferencia durante la polimerización se producía por la presencia del motivo (TGGGGC)₄.

A la vista de los resultados obtenidos, concluimos que los alelos del minisatélite Msh43 que contenían el motivo (TGGGGC)₄ formaban estructuras G-cuadrúplex en presencia de potasio lo que daba lugar a un bloqueo de la síntesis de DNA durante la PCR, provocando la amplificación preferencial del alelo Msh43 73.1, que carecía de ese motivo en su secuencia.

Los resultados que acabamos de describir con el minisatélite Msh43 constituyen una nueva prueba experimental de la formación de estructuras G-cuadrúplex *in vitro*. La capacidad que muestran secuencias obtenidas de DNA genómico para formar dichas estructuras *in vitro* constituye una de las evidencias que sugieren la presencia de G-cuadrúplex *in vivo* (Maizels, 2006). El elevado número de secuencias potencialmente formadoras de G-cuadrúplex que se concentran en regiones del genoma con importantes funciones biológicas, como es el caso de los telómeros, los promotores de genes y algunos puntos calientes de recombinación (Fry, 2007), sugieren que la existencia de estas estructuras en las células está asociada a alguna función biológica. Otra importante evidencia en este sentido viene dada por el descubrimiento de proteínas que interaccionan con los G-cuadrúplex, ya que se han identificado proteínas que promueven su formación (Fang & Cech, 1993), así como también otras que se unen a ellos específicamente (Etzioni et al., 2005), en algunos casos aumentando su estabilidad (Weisman-Shomer & Fry, 1994). Además, también se han descrito proteínas que provocan una desestabilización de los G-cuadrúplex (Harrington et al., 1997), siendo algunas de ellas nucleasas que actúan específicamente sobre dichas estructuras (Sun et al., 2001).

Varias son las funciones biológicas que se ha sugerido podrían desempeñar las estructuras G-cuadrúplex. Entre ellas se encuentra su posible intervención en la regulación de la transcripción génica. Una prueba en esta dirección fue la observación de que una estructura G-cuadrúplex localizada en el promotor de c-MYC actúa como un

elemento represor de la transcripción de este gen (Siddiqui-Jain et al, 2002). También se ha implicado a las estructuras G-cuadrúplex situadas en los telómeros como elementos que bloquean la actividad de la telomerasa, impidiendo la elongación de los mismos (Zahler et al, 1991). Una posible aplicación de esta investigación sería el diseño de una estrategia en la lucha contra el cáncer generando moléculas que estabilicen las estructuras G-cuadrúplex, con el fin de evitar la elongación de los telómeros evitando así la proliferación celular (Neidle & Read, 2000). Además, se ha sugerido que los G-cuadrúplex pueden estar implicados en procesos recombinatorios al demostrarse que dichas estructuras pueden promover recombinación homóloga independiente de RecA en plásmidos (Shukla & Roy, 2006). Incluso ha sido sugerido que los G-cuadrúplex pueden inducir la generación de expansiones de secuencias repetitivas mediante procesos de *slippage*, pudiendo estar relacionados con el origen de desórdenes neurológicos como la epilepsia mioclónica progresiva (Saha & Usdin, 2001). En el caso del minisatélite MsH43, evidentemente todavía no podemos afirmar nada sobre una hipotética función en la célula, si es que la tiene. Futuras investigaciones en nuestro y otros laboratorios podrán arrojar luz sobre si la capacidad de MsH43 de formar estructuras G-cuadrúplex tiene algún papel en la dinámica celular.

Capacidad recombinogénica y G-cuadrúplex

En nuestro laboratorio se han llevado a cabo investigaciones sobre la capacidad recombinogénica de la secuencia minisatélite MsH42 (Boán et al., 1998). Para realizar dichos estudios, se desarrolló un ensayo de recombinación homóloga intramolecular *in vitro* que permite conocer la frecuencia con la que se producen este tipo de eventos recombinatorios en la secuencia objeto de estudio y analizar posteriormente los recombinantes generados en dichos procesos.

Para realizar los ensayos se diseñaron constructos plasmídicos que portaban dos copias de una secuencia minisatélite clonadas en la misma orientación y entre las cuales se encuentra el gen *LacZ*. El plásmido en el que se clonaron dichos insertos fue pBR322, que contiene el gen de resistencia a la ampicilina así como el origen de replicación OriC. El ensayo estaba diseñado de manera que se podían detectar los eventos de recombinación intramolecular catalizados por un extracto proteico nuclear de testículo de rata que tuviesen lugar entre las dos secuencias homólogas clonadas en el plásmido, siguiendo la estrategia que se describe en la Figura 25. Tras una incubación del plásmido con el extracto proteico, si se produce recombinación homóloga intramolecular, el plásmido sustrato generará dos tipos de productos plasmídicos. Uno de ellos contendrá una copia del minisatélite y el gen *LacZ*, pero no el origen de replicación OriC ni el gen de resistencia a la ampicilina, por lo que no se propagará en las bacterias sembradas en placas que contengan dicho antibiótico. El segundo producto plasmídico portará la otra copia del minisatélite, el gen de resistencia a la ampicilina y el origen de replicación OriC, pero no el gen *LacZ*, por lo que generará colonias bacterianas blancas. Esto permite distinguirlas de las colonias transformadas con los plásmidos originales, que al poseer *LacZ*, podrán realizar la α -complementación y formarán colonias azules en presencia de IPTG y X-Gal. Las colonias blancas, que serán por tanto las que presenten los productos de recombinación intramolecular homóloga, se analizarán mediante la digestión del DNA plasmídico con el enzima de restricción adecuado. Durante el proceso de recombinación intramolecular puede producirse un alineamiento perfecto entre ambas secuencias homólogas, dando lugar a un entrecruzamiento igual con lo que se mantiene la organización del minisatélite, generando recombinantes denominados iguales. Por otro lado, puede ocurrir un alineamiento incorrecto de ambos fragmentos de DNA repetitivo, produciéndose un

entrecruzamiento desigual que originará productos de recombinación con variaciones de tamaño en la secuencia minisatélite, es decir, los denominados recombinantes desiguales. Además, también se pueden generar moléculas desiguales mediante

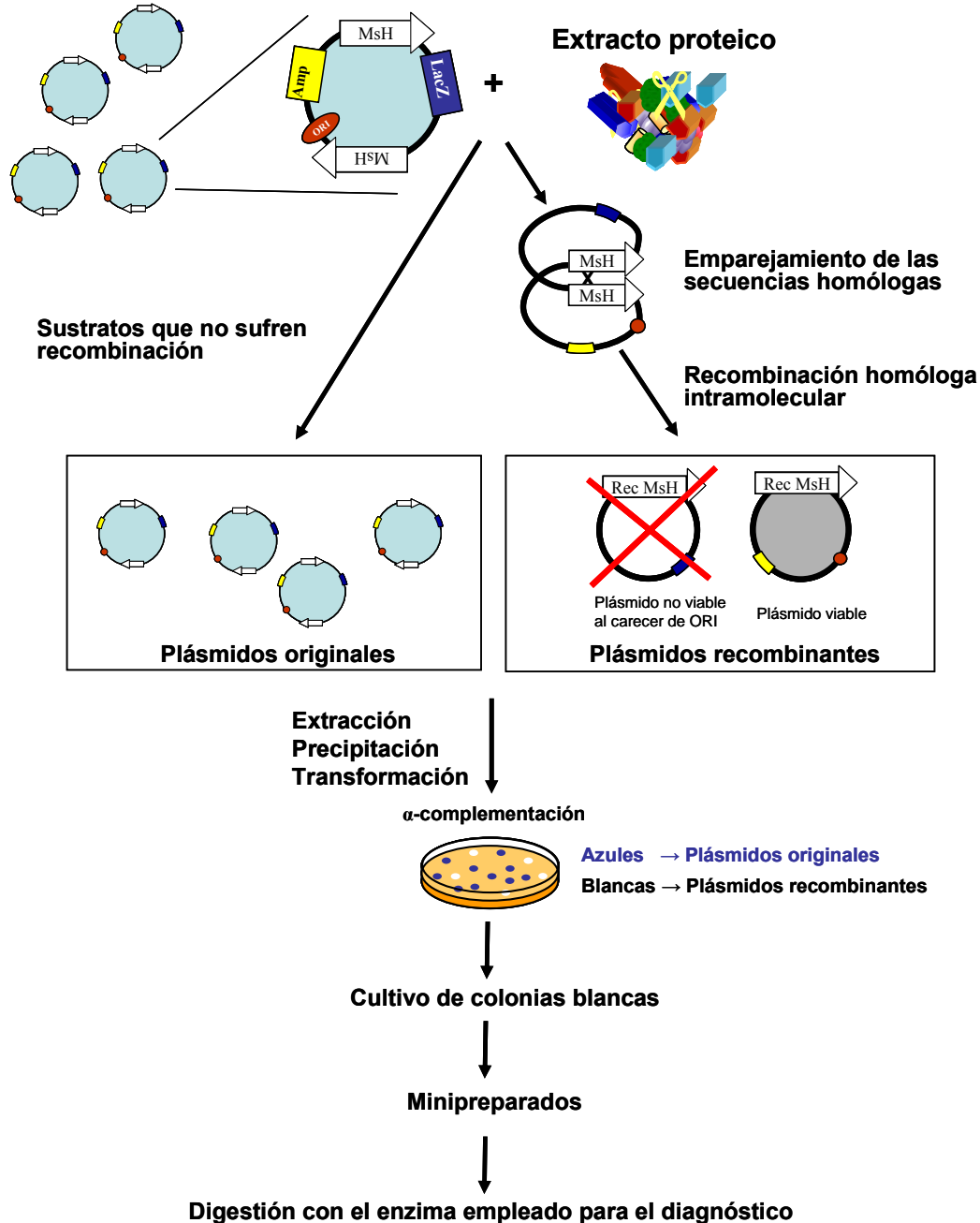


Figura 25. Representación esquemática del ensayo de recombinación homóloga intramolecular. Las flechas blancas Msh representan una secuencia minisatélite y la orientación en la que están clonados en el constructo plasmídico. La caja Amp representa el gen de resistencia a la ampicilina, ORI representa el origen de replicación del plásmido y la caja LacZ el fragmento amino terminal del gen de la β -galactosidasa, que permite el proceso de la α -complementación con las bacterias de genotipo adecuado. La flecha Rec Msh representa la secuencia minisatélite tras haber sufrido un proceso de recombinación homóloga intramolecular.

procesos de *slippage* en la síntesis reparadora de DNA que tienen lugar durante la recombinación, aunque las secuencias se hayan alineado correctamente (Figura 26).

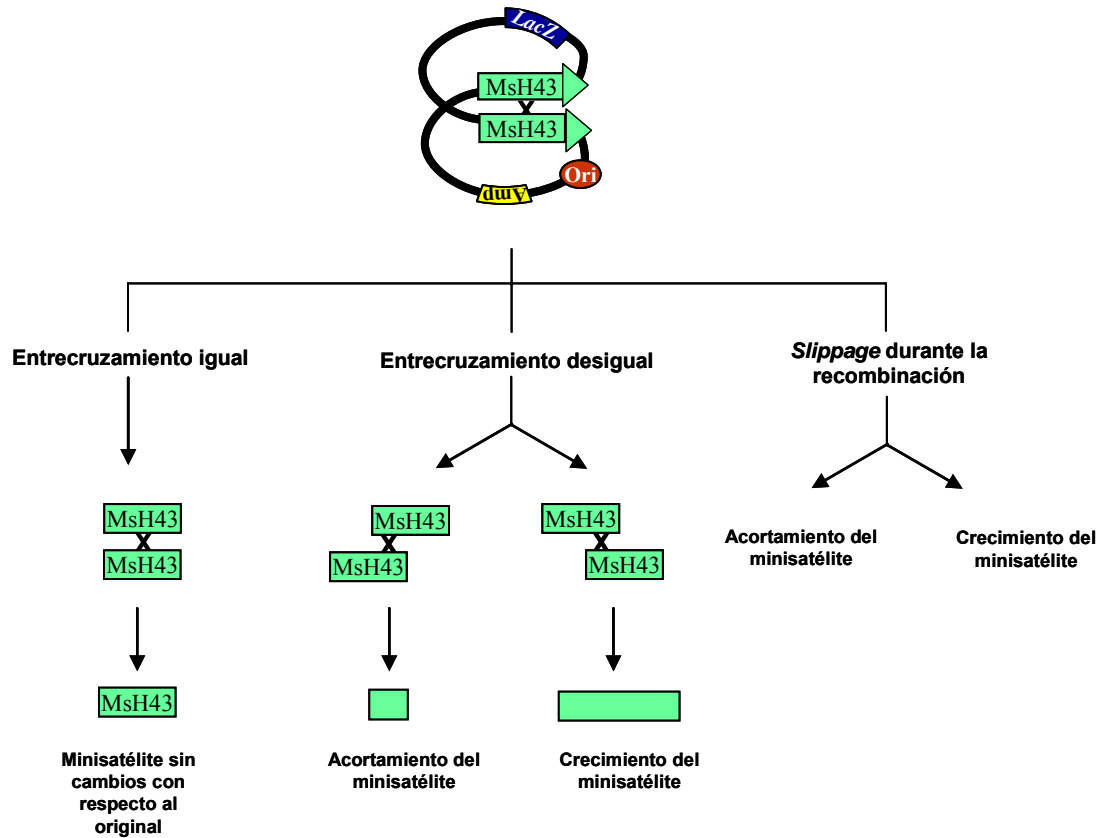


Figura 26. Representación esquemática de los eventos que conducen al mantenimiento o variación de tamaño del minisatélite.

Como se ha descrito previamente en este trabajo, la secuencia minisatélite MsH43 presenta algunos alelos con capacidad de formar estructuras G-cuadrúplex. Este tipo de estructuras se ha demostrado que pueden promover recombinación homóloga independiente de RecA (Shukla & Roy, 2006). Esta relación entre los G-cuadrúplex y los procesos recombinatorios hizo que nos planteáramos investigar si la estructura G-cuadrúplex que describimos en MsH43 puede influir en los procesos recombinatorios que observamos al emplear el ensayo de recombinación homóloga *in vitro* desarrollado en nuestro laboratorio y descrito anteriormente. Para llevar a cabo este estudio se emplearon los alelos MsH43 80.1, que forma G-cuadrúplex, y MsH43 73.1, que no los forma. Para realizar los ensayos se diseñaron los plásmidos p80.1 y p73.1 siguiendo la estrategia descrita previamente. Así, cada plásmido contenía el gen de resistencia a la

ampicilina, portaba dos copias del alelo correspondiente, Msh43 80.1 ó Msh43 73.1, clonadas en la misma orientación y el gen *LacZ* insertado entre ambas copias del minisatélite (Figura 27).

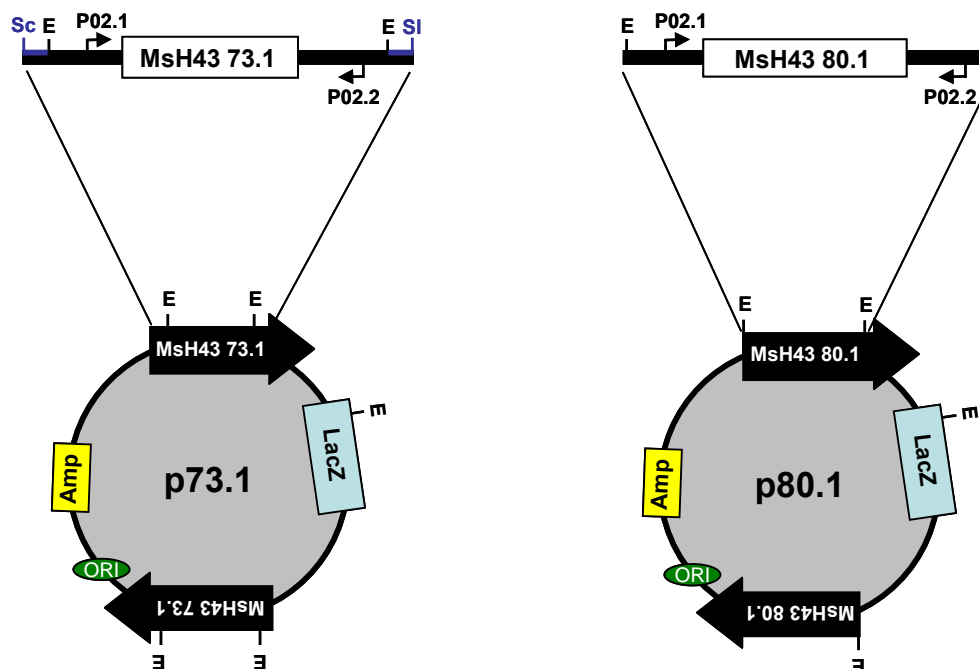


Figura 27. Representación de los plásmidos empleados como sustrato en los ensayos de recombinación homóloga intramolecular in vitro. (A) Esquema del plásmido p73.1. (B) Esquema del plásmido p80.1. Las flechas negras representan la secuencia minisatélite correspondiente y la orientación en la que están clonados en el constructo plasmídico. La caja Amp representa el gen de resistencia a la ampicilina y LacZ al fragmento N-terminal del gen de la β -galactosidasa, que permite el proceso de la α -complementación. P02.1 y P02.2 representan los cebadores empleados en las PCR y que se encuentran situados en las secuencias flanqueantes del minisatélite. E, *EcoRI*; Sc, *SacII*; SI, *SalI*.

Los plásmidos se incubaron con extracto proteico nuclear de testículo de rata como se indica en Materiales y Métodos, realizándose tres réplicas de cada variante del ensayo. Los productos de recombinación generados en estos experimentos, que dan lugar a las colonias blancas, fueron analizados digiriendo su DNA plasmídico con el enzima de restricción *EcoRI*. También se realizó la digestión con *EcoRI* de colonias azules para corroborar que, efectivamente, corresponden a plásmidos originales. En la Figura 28 se muestra un ejemplo del análisis electroforético de los productos de recombinación obtenidos en la digestión de colonias blancas y también los patrones generados por los plásmidos p80.1 y p73.1 originales. Así, la digestión de los productos de recombinación obtenidos de experimentos con p73.1, produce la liberación de un

fragmento que contiene al minisatélite, y que será del mismo tamaño que la secuencia original si deriva de un entrecruzamiento igual y de diferente tamaño si se genera a partir de un entrecruzamiento desigual (Figura 28A). Sin embargo, la digestión de los productos de recombinación obtenidos en los ensayos realizados con p80.1 da lugar a la linearización del plásmido, ya que dicho constructo presenta sólo una diana para *EcoRI* en una de las copias del minisatélite (Figura 28B). Por esta razón, el diagnóstico

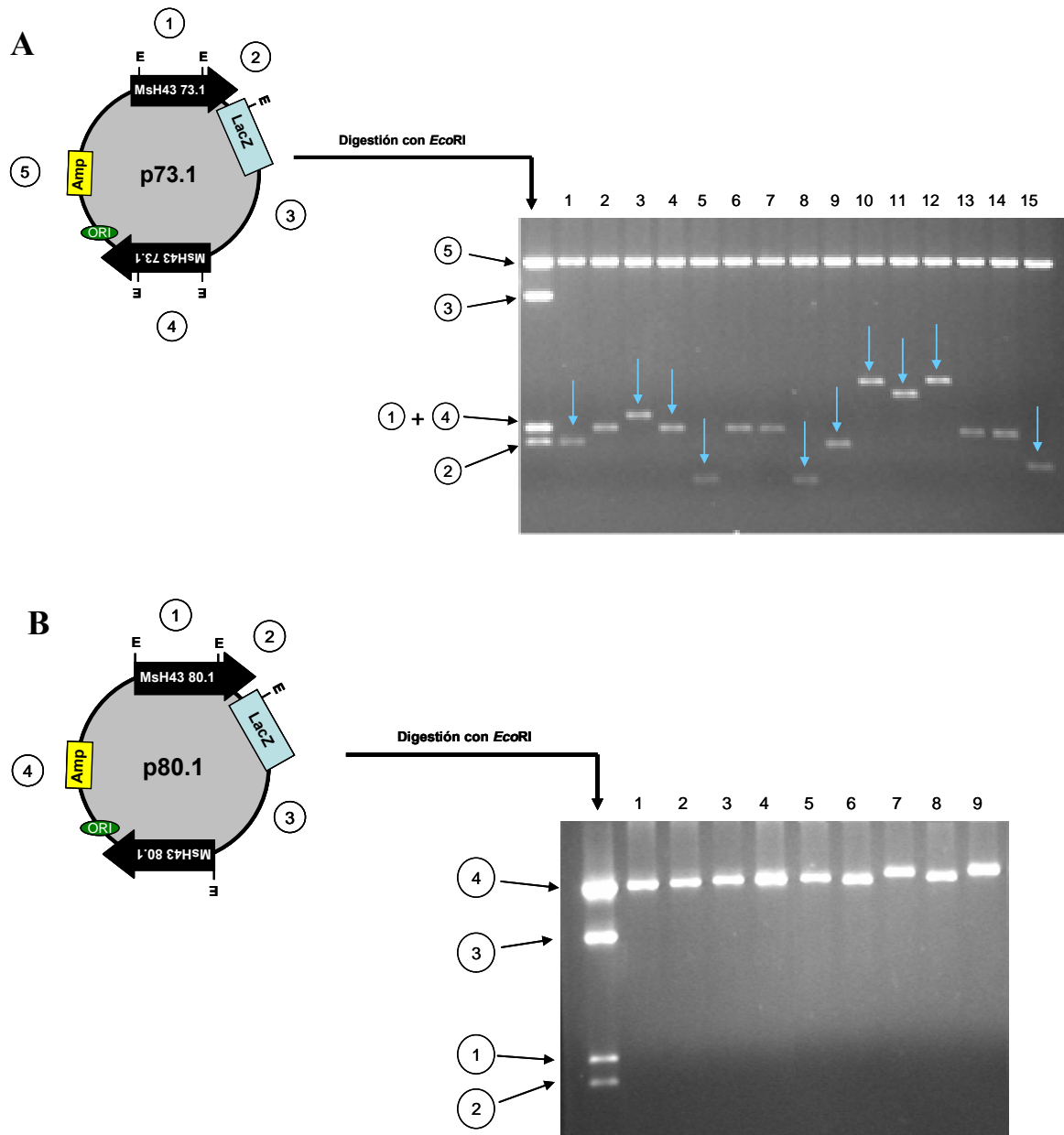


Figura 28. Electroforesis en geles de agarosa al 1,5% que muestra los productos de digestión con *EcoRI* de moléculas recombinantes así como también de los constructos originales. (A) Digestión de recombinantes generados en un ensayo con p73.1 (carriles 1-15). El carril de la izquierda muestra la digestión del plásmido original. Los recombinantes señalados con flechas son aquellos que presentan una secuencia diferente a la original. (B) Digestión de recombinantes obtenidos en un experimento con p80.1 (carriles 1-9). El carril de la izquierda muestra el patrón generado por la digestión del plásmido original.

Enzimático de los ensayos con p80.1 sólo nos permitió distinguir entre plásmidos recombinantes y plásmidos originales. Para posibilitar su discriminación y posterior secuenciación, los productos de recombinación, tanto a los derivados de p80.1 como de p73.1, fueron analizados por PCR empleando los cebadores P02.1/P02.2, de esta forma se consigue amplificar el fragmento de DNA que contiene al minisatélite y parte de las secuencias que lo flanquean. En el caso de p80.1, estas amplificaciones nos permiten realizar una primera diferenciación de los recombinantes iguales y desiguales (Figura 29) y en el caso de p73.1 nos permite confirmar el diagnóstico realizado a partir de la digestión con *EcoRI*.

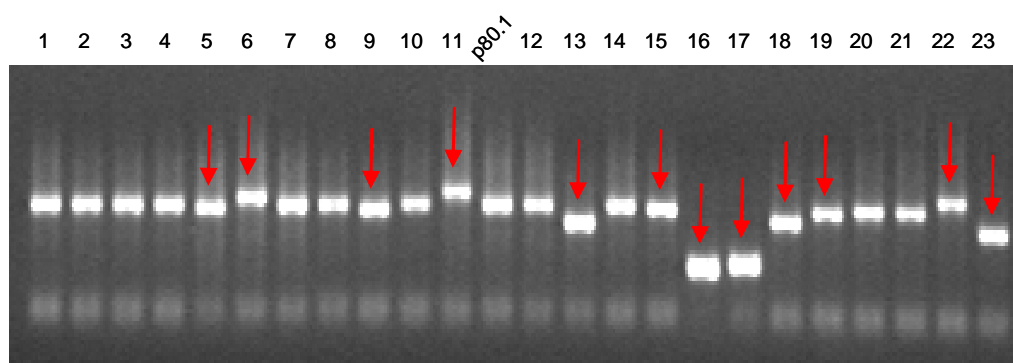


Figura 29. Electroforesis en gel de agarosa al 2% de los productos de la amplificación de plásmidos recombinantes generados a partir de p80.1, realizada con los cebadores P02.1/P02.2 (carriles 1-23). El carril p80.1 indica la amplificación del plásmido original. Las flechas indican las amplificaciones que claramente difieren en tamaño con respecto a la secuencia minisatélite original.

Durante el análisis del polimorfismo de la secuencia minisatélite Msh43, habíamos empleado el método del análisis de heterodúplex para discriminar variantes alélicas de este locus. Cuando se realizaba este ensayo con individuos heterocigotos, siempre se visualizaban en los geles de poliacrilamida moléculas heterodúplex, además de las homodúplex correspondientes. Sin embargo, los individuos homocigotos únicamente generaban la banda homodúplex. La efectividad de esta técnica a la hora de detectar la presencia de moléculas distintas, aunque de tamaño muy similar, la hacía apropiada para verificar que la diferenciación de recombinantes iguales y desiguales que estábamos realizando era correcta. En estos ensayos de heterodúplex se realizaron mezclas del producto de amplificación de los recombinantes con el producto de amplificación del alelo del que procedían, Msh43 80.1 ó Msh43 73.1, y se analizaron en geles de poliacrilamida siguiendo los pasos indicados en material y métodos. En la Figura 30 se presenta una muestra de los análisis de heterodúplex realizados con los

productos de amplificación de recombinantes generados en ensayos con el constructo p80.1. Cuando las mezclas sólo generaban la banda homodúplex indicaba que el recombinante mantenía la secuencia original, mientras que la presencia de moléculas heterodúplex revelaba la presencia de un recombinante desigual. La posterior secuenciación de una muestra de los recombinantes que considerábamos iguales indicó que el diagnóstico que estábamos realizando era correcto y que, efectivamente, estas moléculas conservaban la secuencia original tras sufrir los eventos recombinatorios.

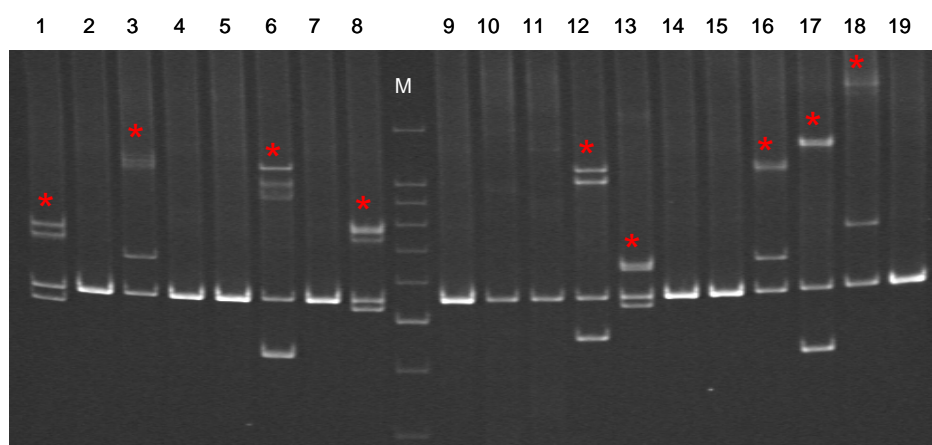


Figura 30. Electroforesis en gel de poliacrilamida al 5% en el que se muestra un análisis de heterodúplex de moléculas recombinantes del plásmido p80.1 (carriles 1-19). El asterisco indica las moléculas heterodúplex que se generan cuando está presente un recombinante desigual. M, 100 bp ladder Promega.

El análisis de los resultados de los ensayos llevados a cabo con p80.1 y p73.1 bajo las condiciones que denominamos estándar, esto es, sin aporte adicional de sal, muestra que la frecuencia de recombinación de ambos plásmidos es similar, 1,072 % para p80.1 y 1,145 % para p73.1 (Tabla 5). Asimismo, la proporción de moléculas recombinantes iguales y desiguales también era semejante en los ensayos realizados con ambos plásmidos. Así, la frecuencia de recombinantes iguales era de 0,442 % y 0,468 % y la de recombinantes desiguales fue de 0,629 % y 0,677 % para los experimentos realizados con p80.1 y p73.1, respectivamente (Tabla 5). Además, cabe destacar que en estos experimentos los eventos recombinatorios que generan productos de recombinación desiguales son más abundantes que los que producen recombinantes que mantienen la secuencia original del minisatélite.

Tabla 5. Resultados obtenidos en los ensayos de recombinación homóloga intramolecular *in vitro*.

Plásmido analizado	Condiciones de ensayo	Nº Colonias <i>LacZ</i> ⁺ / <i>LacZ</i> ⁻	Frecuencia de recombinación (%)	Nº de recombinantes iguales	Frecuencia de recombinantes iguales (%)	Nº de recombinantes desiguales	Frecuencia de recombinantes desiguales (%)
p73.1	Estándar	3863/42	1,087	16	0,414	26	0,673
		4034/45	1,116	21	0,521	24	0,595
		4060/50	1,232	19	0,468	31	0,764
			1,145± 0,077		0,468± 0,054		0,677± 0,085
p73.1	20 mM KCl	8296/39	0,470	24	0,289	15	0,181
		9183/47	0,512	28	0,305	19	0,207
		7850/35	0,448	23	0,293	12	0,155
			0,476± 0,033		0,296± 0,008		0,181± 0,026
p80.1	Estándar	7832/80	1,021	35	0,447	45	0,574
		6095/68	1,116	28	0,459	40	0,657
		6866/74	1,078	29	0,422	45	0,656
			1,072± 0,048		0,442± 0,018		0,629± 0,048
p80.1	20 mM KCl	9450/39	0,413	25	0,265	14	0,148
		12365/47	0,380	35	0,283	12	0,097
		9617/44	0,458	31	0,322	13	0,136
			0,417± 0,039		0,290± 0,029		0,127± 0,027
p80.1	20 mM NaCl	19559/29	0,148	23	0,118	6	0,030
		17360/27	0,156	21	0,121	6	0,035
		16261/21	0,129	16	0,098	5	0,031
			0,144± 0,014		0,112± 0,013		0,032± 0,003
p80.1	20 mM NH ₄ Cl	8114/23	0,283	18	0,222	5	0,061
		8555/25	0,292	21	0,245	4	0,047
		7962/22	0,276	17	0,314	5	0,062
			0,284± 0,008		0,260± 0,048		0,057± 0,008

La frecuencia de recombinación expresa la relación entre las colonias *LacZ*⁺ y las colonias totales para cada plásmido y experimento. La frecuencia de recombinantes iguales y desiguales expresa el porcentaje de entrecruzamientos iguales y desiguales, respectivamente, respecto del total de entrecruzamientos analizados en cada experimento para cada tipo de ensayo y plásmido.

Ya que bajo condiciones estándar no apreciamos diferencias en el comportamiento de p80.1 y p73.1, se realizaron ensayos con un aporte adicional de potasio en las reacciones de recombinación con el fin de favorecer la estabilización de los G-cuadrúplex que pudiese formar el motivo (TGGGGC)₄ del alelo Msh43 80.1. Para determinar la concentración de potasio adecuada se realizaron ensayos con reacciones que contenían concentraciones de 100 mM, 50 mM y 20 mM de KCl. El análisis de los resultados de estos experimentos puso de manifiesto que el número de recombinantes generados disminuye a medida que aumenta la concentración de KCl en la reacción. Se determinó que la concentración de KCl apropiada para llevar a cabo los

ensayos de recombinación era de 20 mM, ya que era la máxima cantidad de KCl a la que se podía realizar los experimentos sin que la frecuencia de recombinación disminuyese de manera importante. Mientras que en los ensayos en condiciones estándar la frecuencia de recombinación de p80.1 y p73.1 fue de 1,072% y 1,145% respectivamente, en los ensayos realizados con aporte adicional de KCl la frecuencia de recombinación fue de 0,417% para p80.1 y 0,476% para p73.1. Este descenso de la frecuencia de recombinación podría explicarse por la disminución de la actividad nucleásica que se observa a medida que se incrementa la concentración de KCl en la reacción (Figura 31). Estos resultados contrastan con las investigaciones realizadas por

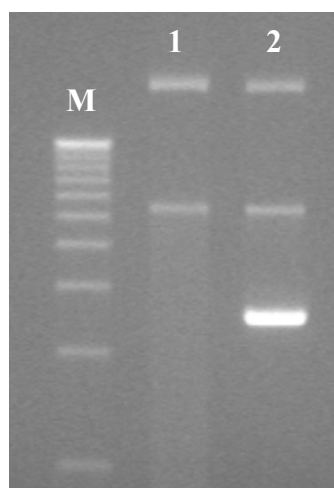


Figura 31. Análisis electroforético de dos reacciones de recombinación correspondientes a un ensayo en condiciones estándar (carril 1) y a uno con aporte de 20 mM KCl (carril 2). M, 100 bp ladder Promega.

Liu et al. (2004) en las que se observa que la actividad recombinasa de hRAD51, proteína implicada en los eventos de recombinación homóloga, es estimulada *in vitro* por la presencia de una concentración salina elevada. El análisis de los datos obtenidos en nuestros ensayos demostró que las frecuencias de recombinación de ambos plásmidos, p73.1 y p80.1, así como la proporción de recombinantes iguales y desiguales entre ellos eran similares, al igual que había sucedido en los experimentos realizados en condiciones estándar (Tabla 5). Por lo tanto, parece claro que la presencia de las estructuras G-cuadrúplex no influye en los procesos de recombinación que se detectan con nuestro sistema de recombinación homóloga intramolecular *in vitro* o que no se generan o estabilizan de manera suficiente como para ser detectadas.

Un dato muy interesante fue que al comparar los resultados obtenidos en los ensayos estándar con aquellos que llevan 20 mM de KCl, se observa que la proporción de recombinantes iguales en los experimentos con aporte de esta sal aumenta notablemente, utilizando indistintamente p80.1 ó p73.1. Este ascenso se debió fundamentalmente a un importante descenso en la frecuencia de recombinantes desiguales con respecto a la de los recombinantes iguales (Tabla 6). En términos porcentuales, la disminución de las moléculas desiguales fue del 79,936% para p80.1 y 73,373% para p73.1, mientras que la de moléculas iguales fue del 34,459% para p80.1 y 36,887% para p73.1 (Tabla 6). La observación de este efecto con ambos plásmidos sugiere que no está relacionado con la capacidad de formación de G-cuadrúplex, ya que el alelo Msh43 73.1 presente en el constructo p73.1 no los forma.

El hecho de que en los ensayos realizados en condiciones estándar se obtenga una mayor proporción de moléculas desiguales, mientras que en los experimentos con aporte de KCl la mayoría de los recombinantes sean iguales, podría ser debido tanto a la presencia de KCl como al aumento de la fuerza iónica provocada por el incremento en la concentración de dicha sal. Para resolver esta cuestión, se realizaron ensayos de recombinación en los que se adicionaron otras sales, concretamente NaCl y NH₄Cl, manteniendo la misma fuerza iónica que en los experimentos con aporte de KCl. Puesto que el comportamiento de p80.1 y p73.1 en los ensayos de recombinación había demostrado ser igual, los experimentos con NaCl o NH₄Cl se llevaron a cabo sólo con p80.1. El análisis de los resultados de estos ensayos mostró que, al igual que ocurría cuando se aportaba KCl, existe un pronunciado descenso de la frecuencia de recombinación y de la actividad nucleásica con respecto a la que presentan los ensayos en condiciones estándar. Así, mientras que en condiciones estándar la frecuencia de recombinación obtenida era de 1,072%, en los experimentos con NaCl y NH₄Cl fue de 0,144% y 0,284%, respectivamente. En otras palabras, la frecuencia de recombinación descende un 68,106% en los ensayos con NaCl y 65,47% en los experimentos con NH₄Cl, con respecto a la obtenida en experimentos con KCl. En nuestro estudio, por lo tanto, no detectamos ninguna relación entre la frecuencia de recombinación y el tamaño del catión presente. Por el contrario, otras investigaciones han mostrado una relación entre el tamaño del catión presente en una reacción y la influencia sobre la actividad de hRAD51, una proteína que como se mencionó anteriormente está implicada en procesos recombinatorios (Shim et al., 2006).

Tabla 6. Descenso de las frecuencias de recombinación obtenidas en los ensayos con aporte de diferentes sales con respecto a las frecuencias obtenidas en condiciones estándar

Plásmido analizado	Condiciones del ensayo	Descenso de la frecuencia de recombinantes iguales (%)	Descenso de la frecuencia de recombinantes desiguales (%)
p73.1	20 mM KCl	36,887	73,373
p80.1	20 mM KCl	34,459	79,936
p80.1	20 mM NaCl	74,775	94,904
p80.1	20 mM NH₄Cl	48,874	90,924

El descenso en las frecuencias de recombinación expresa la relación entre los recombinantes obtenidos con ambos plásmidos en experimentos estándar y con aporte de distintas sales.

En cuanto a la generación de recombinantes iguales y desiguales en los ensayos realizados con aporte de NaCl o NH₄Cl, al igual que sucedía en los experimentos con KCl, observamos un claro sesgo hacia una mayor proporción de moléculas iguales. La causa de este sesgo fue, al igual que ocurría con la adición de KCl, el pronunciado descenso de la frecuencia de recombinantes desiguales en estos ensayos, que fue del 90,924% para NH₄Cl y 94,904% para NaCl, mientras que la frecuencia de recombinantes iguales desciende un 48,874% para NH₄Cl y un 74,775% para NaCl, al compararlas con las obtenidas en condiciones estándar (Tabla 6).

Los resultados obtenidos indican, por tanto, que la disminución en la frecuencia de recombinación así como el aumento de la proporción de recombinantes que mantienen la secuencia original está relacionada con un incremento de la fuerza iónica en la reacción más que con la presencia de un ión en concreto. Una hipótesis para explicar estos resultados sería que el aumento de la fuerza iónica produjera cambios conformacionales en las proteínas implicadas en los eventos recombinatorios que tienen

lugar en nuestro sistema, de manera que estas proteínas resolviesen de manera más eficiente los entrecruzamientos iguales que los desiguales. Efectos en esta dirección han sido descritos por otros investigadores (Liu et al., 2004). Otra posible explicación se basa en el hecho de que cuando en los ensayos existe un aporte de alguna sal la actividad nucleásica disminuye, por lo que sería posible que al haber un menor número de moléculas que requieran ser reparadas, la maquinaria recombinogénica sea más eficiente.

Para conocer dónde se estaban produciendo las reorganizaciones en los recombinantes desiguales con respecto a los dos alelos originales de Msh43 empleados en los ensayos así como el grado de complejidad de las mismas, se realizó un análisis de la secuencia de una muestra de estas moléculas (Figuras 32-35). Este estudio mostró la conservación de las secuencias que flanquean al minisatélite, las cuales no parecen estar implicadas en las reorganizaciones que tienen lugar. De los 175 recombinantes analizados sólo en tres se ve involucrado uno de los extremos de la secuencia repetitiva, pero la reorganización que se produce sólo abarca al minisatélite.

La mayoría de los recombinantes presentan reorganizaciones en las que se ha producido una duplicación o delección simple, a excepción de las moléculas S5 y S11, generadas en ensayos con p80.1 bajo condiciones estándar, y K10 y K13, procedentes de ensayos con p80.1 y con aporte adicional de KCl, que han sufrido la delección de dos fragmentos de la secuencia con respecto a la molécula original (Figura 32). Además, las duplicaciones son directas exceptuando el recombinante S5 generado en un ensayo en condiciones estándar con p73.1, en el que hay una unidad repetitiva entre el fragmento duplicado y su correspondiente copia (Figura 33). En nuestro análisis las reorganizaciones tampoco dan lugar a la aparición de unidades repetitivas truncadas, conservándose la organización de motivos en tándem característica de la secuencia minisatélite. Estos resultados contrastan con los obtenidos por Berg et al. (2003), ya que al analizar los recombinantes del minisatélite humano MS1 obtenidos del estudio de DNA genómico procedente de células germinales observan que la mayoría de las reorganizaciones generadas son complejas, con múltiples duplicaciones imperfectas de bloques de unidades repetitivas, dando lugar en muchos casos a motivos truncados. En esta dirección, Lopes et al. (2006) al estudiar las moléculas obtenidas en los procesos recombinatorios sobre el minisatélite CEB1 insertado en *Saccharomyces cerevisiae*, también muestra que la mayoría de las moléculas generadas presentan reorganizaciones complejas, sin embargo no detecta la aparición de unidades repetitivas truncadas.

En cuanto al número de ocasiones en las que detectamos cada reorganización, entre las moléculas generadas a partir de p73.1 sólo encontramos una reorganización repetida: S12, recombinante procedente de un ensayo en condiciones estándar y de mayor tamaño que la secuencia original (Figura 33). Sin embargo, en los recombinantes generados a partir de p80.1 detectamos nueve reorganizaciones que han sido observadas en más de una ocasión. Entre estos se encuentran moléculas que presentan duplicaciones con respecto a la secuencia de la que proceden: recombinantes S1, S9, S12 y S13 (procedentes de ensayos en condiciones estándar), K3 y K7 (experimentos realizados con 20 mM KCl) y N6 (experimentos realizados con 20 mM NH₄Cl) (Figura 34); y deleciones: S6 (generado en condiciones estándar) y K14 (procedente de ensayo con aporte de KCl) (Figura 32). La existencia en nuestro análisis de reorganizaciones que aparecen en más de una ocasión contrasta con los resultados obtenidos por Lopes et al. (2006), puesto que en su estudio los recombinantes son únicos, de modo que cada reorganización aparece una sola vez, aunque esto puede deberse al menor número de muestras analizadas.

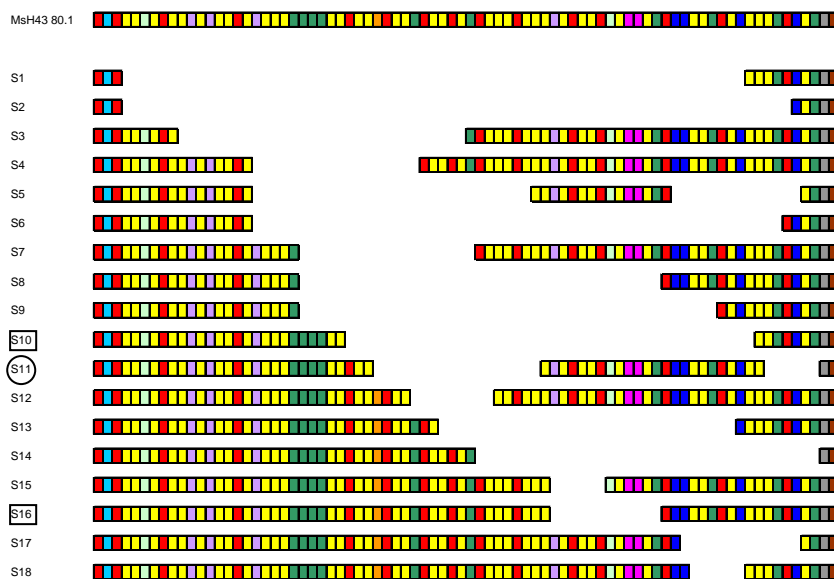
Por otra parte, en nuestro análisis no se aprecian variaciones en las secuencias recombinantes que se puedan deber a eventos mutagénicos puntuales, a excepción del recombinante S1, de mayor tamaño que la secuencia original y generado en ensayos con p73.1 bajo condiciones estándar, en el que se produce la mutación del motivo 30 y una inserción de dos unidades repetitivas entre los motivos 35 y 36 (Figura 33).

Los recombinantes que presentan duplicaciones con respecto a la secuencia original y que fueron generados en ensayos con p80.1 y con aporte de KCl, presentan una característica interesante, en ninguno se ve implicada la secuencia que se encuentra a 5' del motivo (TGGGGC)₄ en las reorganizaciones. Sin embargo, en los ensayos realizados también con p80.1 pero en condiciones estándar o con NaCl o NH₄Cl, sí hay recombinantes que presentan duplicaciones en esa zona. Al analizar las moléculas obtenidas en ensayos con p73.1, se observa que entre ellas existen algunas en las que las duplicaciones afectan a zonas situadas a 5' del motivo (TGGGGC)₃. En cuanto a los recombinantes que presentan deleciones con respecto a la molécula original, ya sea p80.1 ó p73.1, e independientemente de las condiciones del ensayo, existen moléculas en las que el fragmento delecionado está situado a 5' de la zona en la que se localizan las repeticiones de TGGGGC.

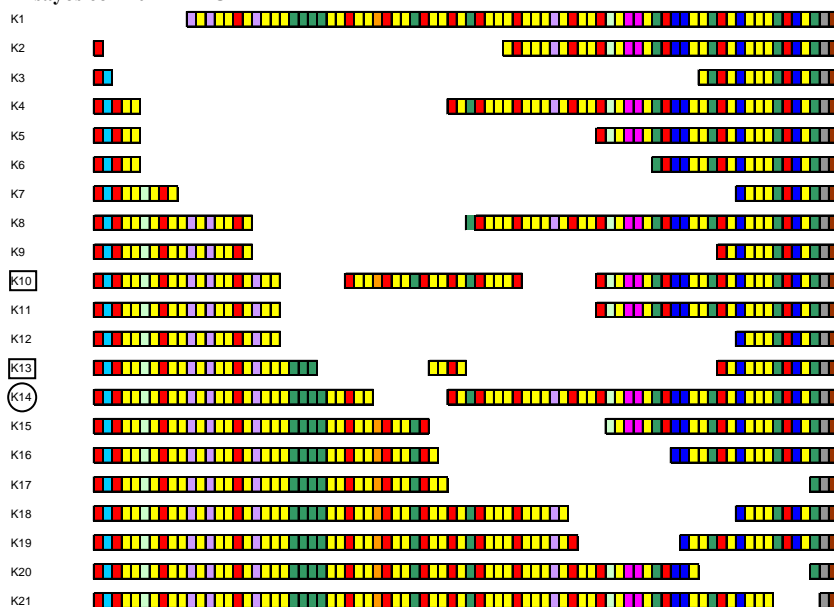
El hecho de que el sesgo sea específico de p80.1, que únicamente se observe en los productos de ensayos realizados con aporte adicional de potasio y que en la posición

que limita la zona donde se producen las duplicaciones se encuentre el motivo (TGGGGC)₄, sugieren que este efecto podría estar relacionado con la presencia de estructuras G-cuadrúplex.

Ensayos estándar



Ensayos con 20 mM KCl



Ensayos con 20 mM NaCl

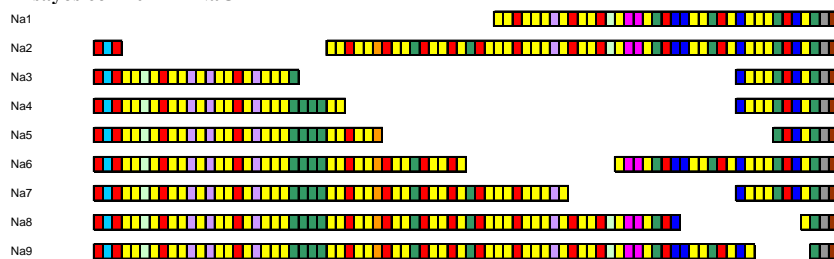
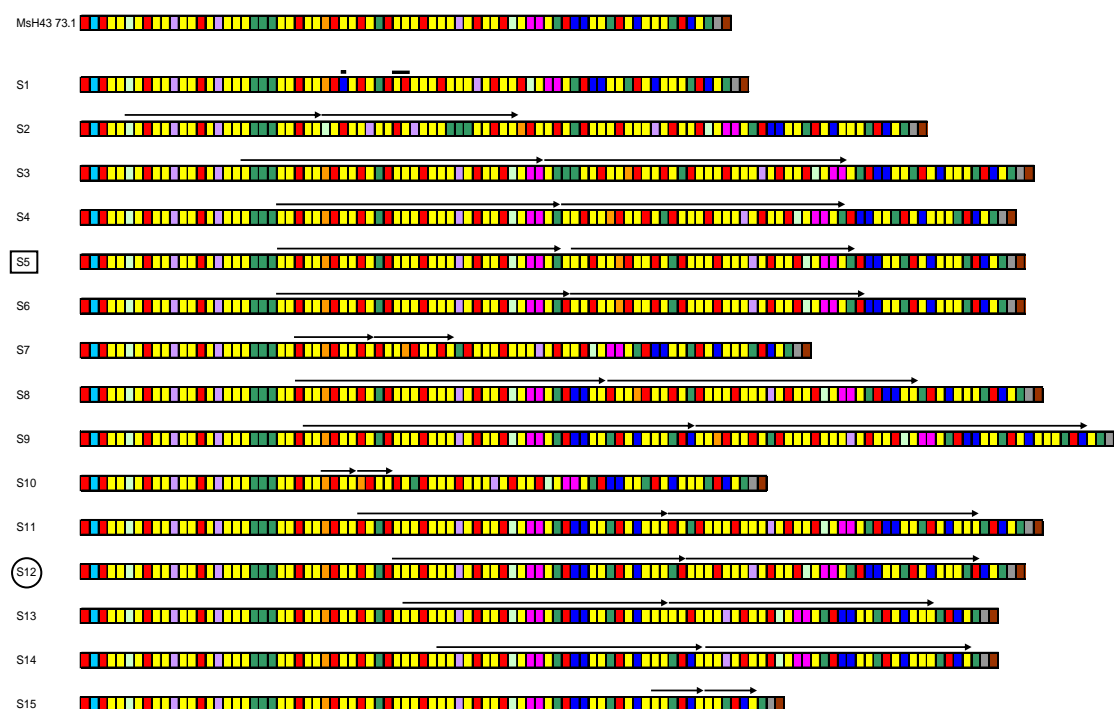
Ensayos con 20 mM NH₄Cl

Figura 32. Secuencias de algunos de los recombinantes obtenidos en los ensayos de recombinación bajo condiciones estándar o con aporte de KCl, NaCl o NH₄Cl, en los que el minisatélite sufre un acortamiento. La discontinuidad en la secuencia indica el fragmento delecionado. Los recombinantes que presentan la deleción de dos fragmentos de la secuencia con respecto a la molécula original están señalados con un rectángulo y los que aparecen en más de una ocasión están indicados con un círculo. En la parte superior se muestra la organización del alelo original MsH43 80.1.

Ensayos estándar



Ensayos con 20 mM KCl

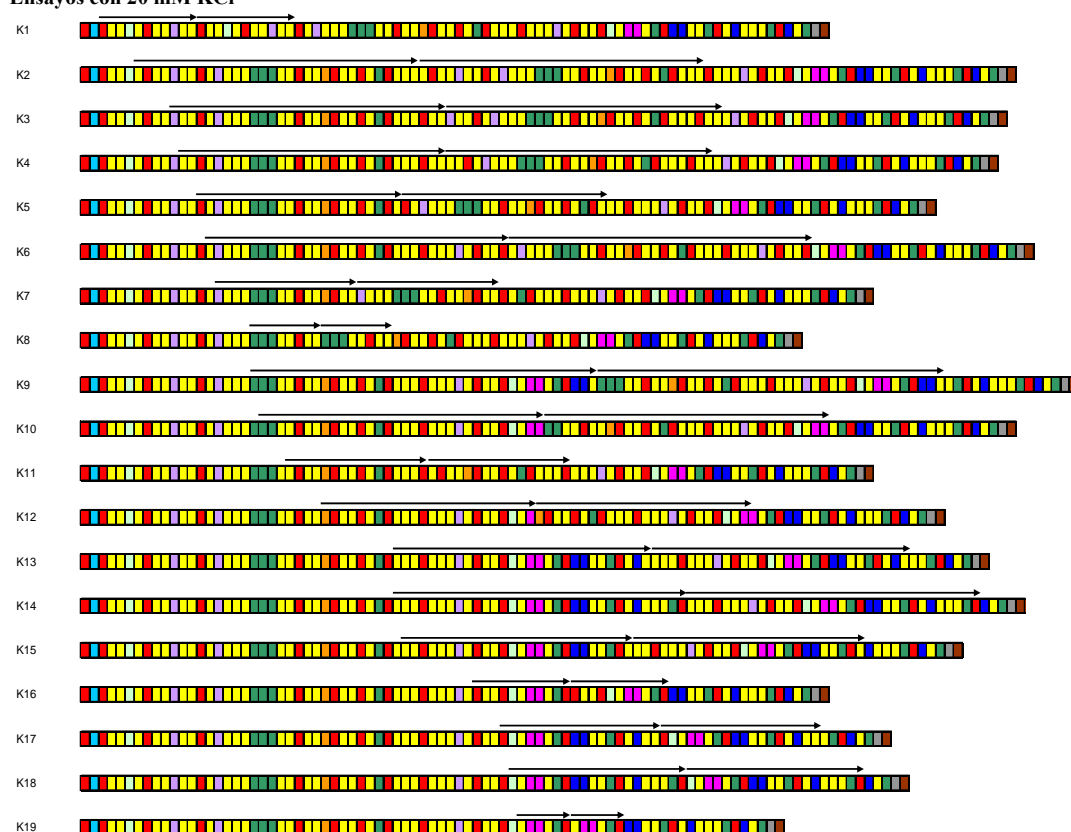
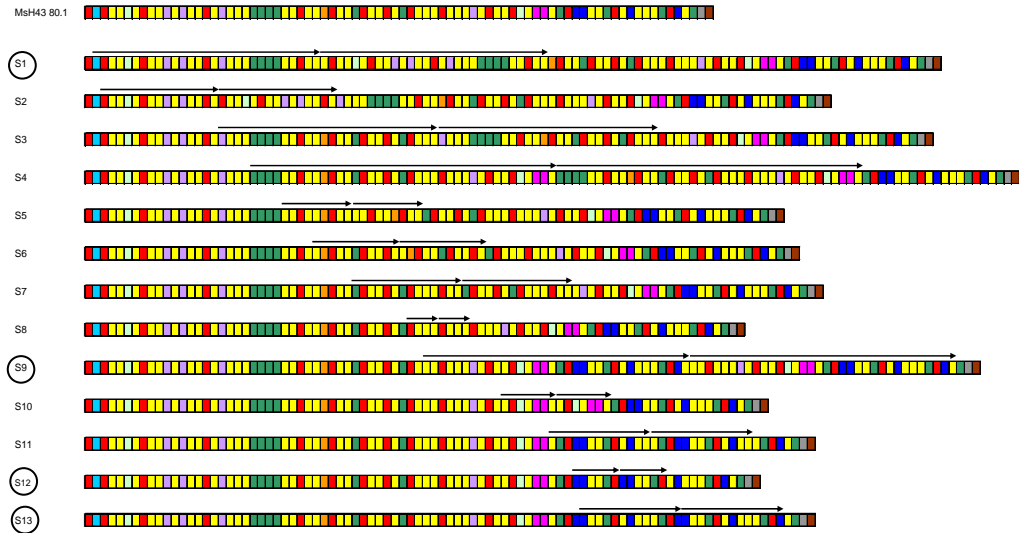


Figura 33. Secuencias de algunos de los recombinantes obtenidos en los ensayos de recombinación bajo condiciones estándar o con aporte de KCl, en los que el minisatélite sufre un alargamiento. Las flechas indican la secuencia duplicada. Los recombinantes que presentan la delección de dos fragmentos de la secuencia con respecto a la molécula original están señalados con un rectángulo y los que aparecen en más de una ocasión están indicados con un círculo. En el recombinante S1 el punto indica un motivo mutado y al guión la inserción de las dos unidades repetitivas. En la parte superior se muestra la organización del alelo original MsH43 73.1.

Ensayos estándar



Ensayos con 20 mM KCl



Ensayos con 20 mM NaCl

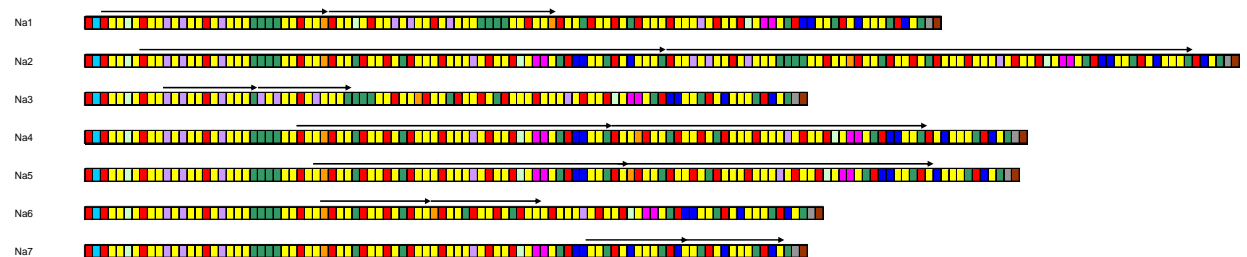
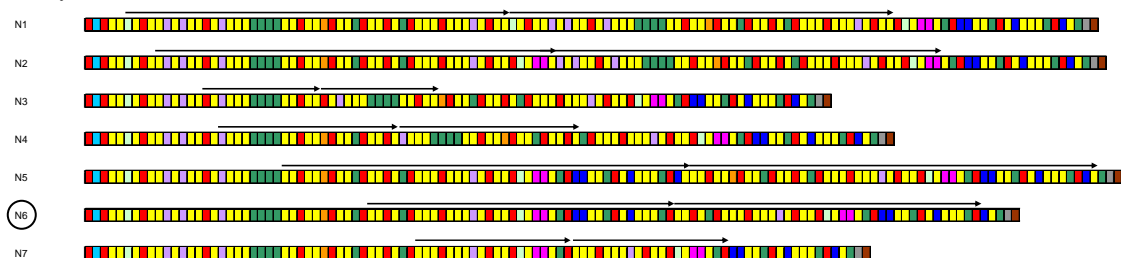
Ensayos con 20 mM NH₄Cl

Figura 34. Secuencias de algunos de los recombinantes obtenidos en los ensayos de recombinación bajo condiciones estándar o con aporte de KCl, NaCl o NH₄Cl, en los que el minisatélite sufre un alargamiento. Las flechas indican la secuencia duplicada. Los recombinantes que presentan la delección de dos fragmentos de la secuencia con respecto a la molécula original están señalados con un rectángulo. En la parte superior se muestra la organización del alelo original MsH43 80.1.

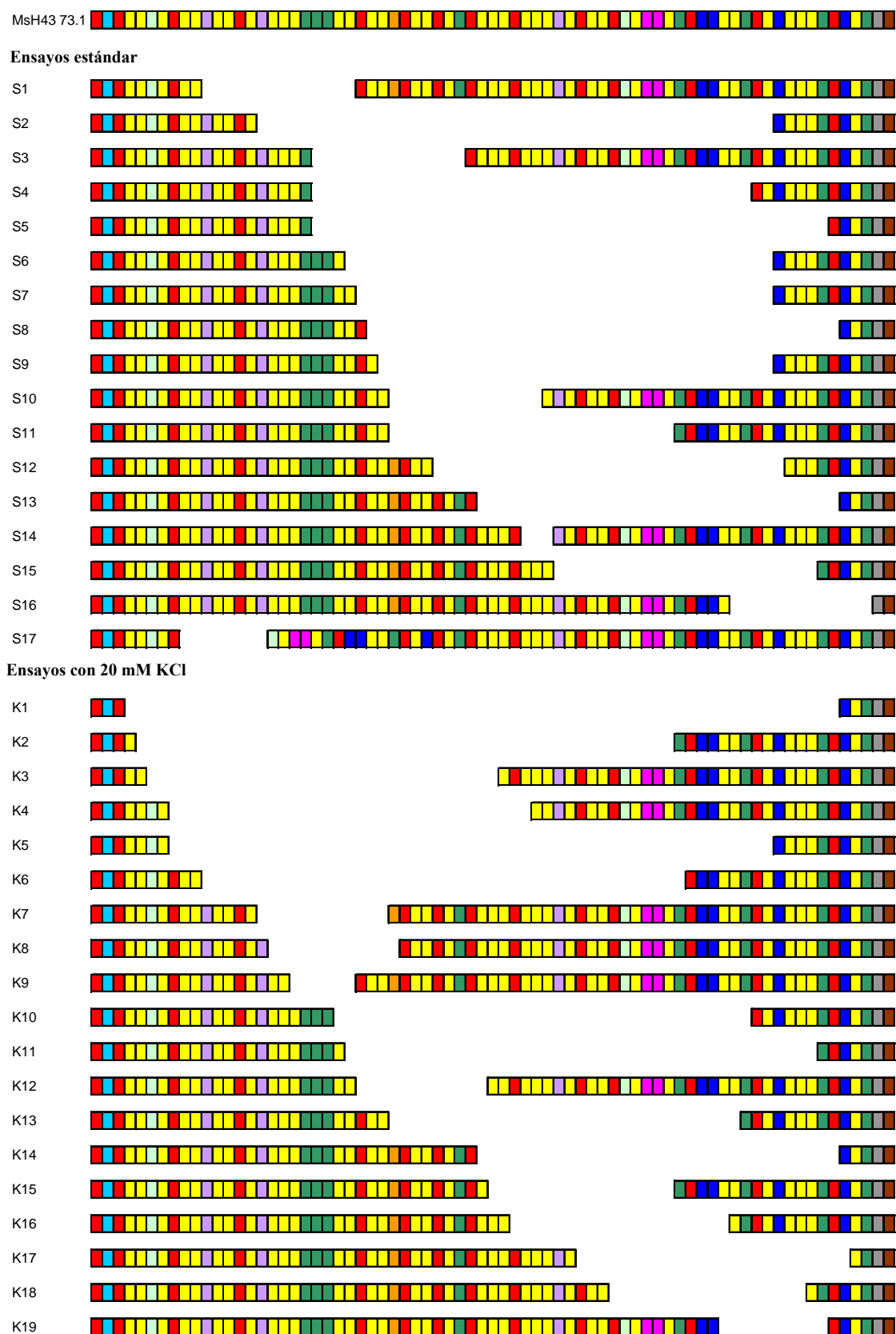


Figura 35. Secuencias de algunos de los recombinantes obtenidos en los ensayos de recombinación bajo condiciones estándar o con aporte de KCl, en los que el minisatélite sufre un acortamiento. La discontinuidad de la secuencia indican el fragmento deleccionado. En la parte superior se muestra la organización del alelo original Msh43 73.1.

Nacimiento e historia evolutiva de MsH43

Una forma de abordar el análisis de la dinámica de un genoma es estudiar la evolución de cada una de las pequeñas partes que lo integran (Pääbo, 2003). En este sentido, muchas secuencias de DNA repetitivo poseen unos mecanismos evolutivos característicos que difieren de los de las secuencias de copia única (Charlesworth et al., 1994). Puesto que en muchos organismos el DNA repetitivo constituye una fracción sustancial del genoma, la comprensión de los procesos que gobiernan su evolución puede ayudarnos a entender cómo se ha construido y desarrollado nuestro genoma hasta la actualidad. Existen diversos estudios sobre la evolución de secuencias repetitivas, como los miembros de la familia *Alu* o secuencias compuestas de repeticiones de dinucleótidos. Por su parte, la mayoría de los estudios realizados con micro- y minisatélites se han llevado a cabo empleando secuencias constituidas por unidades repetitivas perfectas (Almeida & Penha-Gonçalves, 2004), siendo la historia evolutiva de los micro- y minisatélite complejos, entre los que se encuentra la secuencia MsH43, menos estudiada. Cabe señalar, además, que los datos disponibles sobre la evolución de secuencias repetitivas provienen principalmente de estudios intraespecíficos o mediante la comparación de especies muy próximas, que representan una zona limitada del árbol filogenético (Zhu et al., 2000; Webster et al., 2002).

El análisis del polimorfismo de MsH43 en humanos dio lugar, tal como describimos con anterioridad, a la identificación de ocho alelos, presentando entre 71-83 unidades repetitivas de 5-6 pb, habiéndose identificado 11 motivos diferentes. Utilizando las bases de datos disponibles, se realizó la búsqueda de secuencias homólogas a la secuencia humana MsH43 en los genomas de los siguientes mamíferos: perro, gato, vaca, ratón, rata, cerdo, chimpancé y macaco rhesus. Los datos obtenidos mostraron que solamente en los genomas de los dos primates se hallaba una secuencia repetida en tándem ortóloga a MsH43, lo que indica que dicha secuencia es exclusiva de este orden. Sin embargo, en el genoma del perro encontramos secuencias homólogas a aquellas que flanquean MsH43 en el genoma de humano, de chimpancé y de macaco rhesus.

Para analizar la organización de este locus en el orden de los primates, la secuencia MsH43 fue amplificada por PCR a partir de individuos de diferentes especies de primates representativas de las principales ramas de su árbol filogenético (Figura 36). Los cebadores P1 y P02.2, diseñados a partir de la secuencia humana, sólo produjeron

productos de amplificación en gibón, orangután, gorila y chimpancé. Para amplificar el locus Msh43 de las otras especies, se diseñó el cebador P2.1 utilizando la secuencia del locus de macaco rhesus obtenida de las bases de datos genómicas. Con la combinación P1/P2.1, se obtuvieron diversos productos de PCR a partir de distintas especies (macaco, mandril y mangabey) de Monos del Viejo Mundo (MVM). Finalmente, para amplificar el DNA genómico de Monos del Nuevo Mundo (MNM) y Prosimios, la estrategia empleada consistió en diseñar un nuevo grupo de cebadores (P2.2, P2.3, P3 y P3.1) a partir de las regiones conservadas que se detectan en un alineamiento de las secuencias flanqueantes de *Macaca mulata* (un MVM) y la secuencia ortóloga en perro. La combinación P3/P2.2 generó un producto de amplificación en ambas especies de MNM, capuchino y tamarindo, mientras que en lemur sólo se obtuvo empleando P3.1/P2.3.

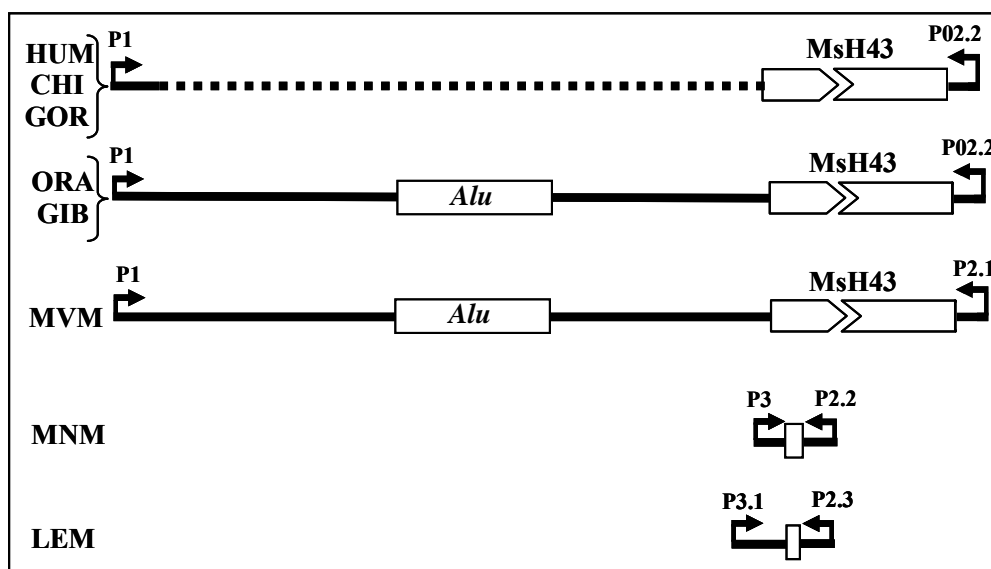
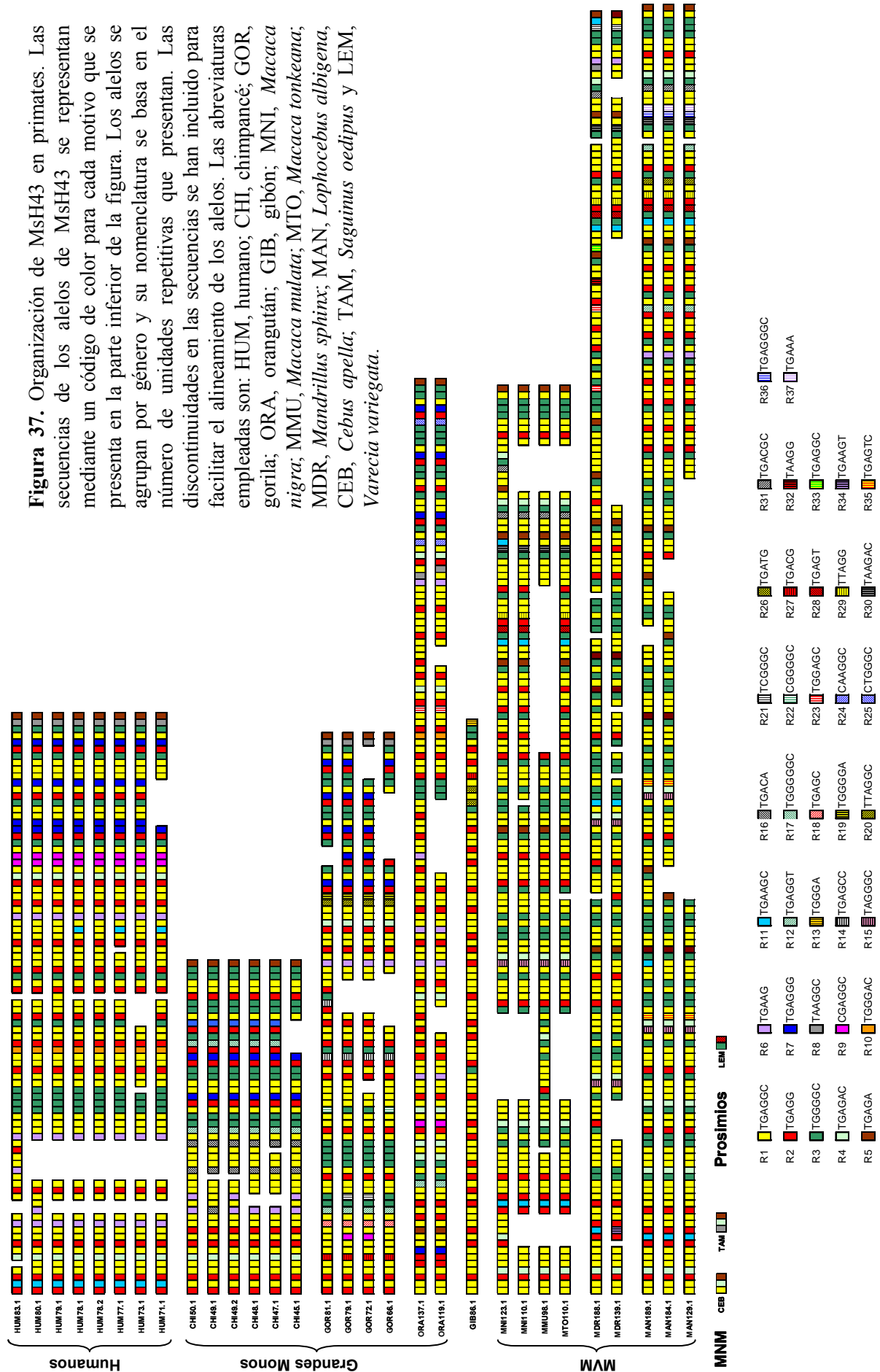


Figura 36. Representación esquemática de los productos de amplificación del locus Msh43 de distintas especies de primates. La discontinuidad en la caja Msh43 simboliza los diferentes tamaños alélicos y la línea punteada indica la delección del fragmento de DNA que incluye la secuencia *Alu*. La posición de los cebadores se señala con flechas. Las abreviaturas empleadas son: HUM, humano; CHI, chimpancé; GOR, gorila; ORA, orangután; GIB, gibón; MVM, monos del viejo mundo; MNM, monos del nuevo mundo y LEM, lemur. Los tamaños de los productos de PCR son: HUM, 603-672 pb; CHI, 470-500 pb; GOR, 590-670 pb; ORA, 2.3-2.4 kb; GIB, 2.1 kb; MVM, 2.3-2.7 kb; MNM, 220 pb y LEM, 430 pb.

Los diferentes productos de amplificación obtenidos fueron secuenciados, realizándose una comparación intra- e interespecífica de las secuencias. Para ello, al igual que se había hecho anteriormente para analizar los distintos alelos humanos, se alinearon dichas secuencias empleando un código de color para cada unidad repetitiva diferente (Figura 37).

Figura 37. Organización de MsH43 en primates. Las secuencias de los alelos de MsH43 se representan mediante un código de color para cada motivo que se presenta en la parte inferior de la figura. Los alelos se agrupan por género y su nomenclatura se basa en el número de unidades repetitivas que presentan. Las discontinuidades en las secuencias se han incluido para facilitar el alineamiento de los alelos. Las abreviaturas empleadas son: HUM, humano; CHI, chimpancé; GOR, gorila; ORA, orangután; GIB, gibón; MNI, *Macaca nigra*; MMU, *Macaca mulata*; MTO, *Macaca tonkeana*; MDR, *Mandrillus sphinx*; MAN, *Lophocebus albigena*; CEB, *Cebus apella*; TAM, *Saguinus oedipus* y LEM, *Varecia variegata*.



El análisis de los resultados muestra que los Prosimios y MNM presentan una secuencia corta compuesta por dos y tres motivos, respectivamente, que no forman una organización en tándem. En los MVM aparece por primera vez un fragmento largo de DNA con una organización repetida en tándem homóloga al Msh43 humano. La comparación de las secuencias de Msh43 de las distintas especies de MVM analizadas (macaco, mandril y mangabey) mostraban un fragmento de 18 unidades repetitivas bien conservadas en el extremo 5', mientras que en la región del extremo 3' hay 15 motivos idénticos entre mangabey y macaco pero diferente en mandril. La zona interna de Msh43 presenta varias duplicaciones y algunos motivos nuevos. En gibón, Msh43 presenta una organización peculiar ya que muestra una notable homología en el contenido de unidades repetitivas. Así, en esta especie los motivos R1 y R2 representan ~ 90.7% de las unidades repetitivas totales, mientras que en MVM representan ~ 63.5%. En orangután la estructura de Msh43 sólo mantiene los tres primeros motivos del extremo 5' de MVM y gibón y no se observa una similitud destacable con respecto al extremo 3' de MVM, a excepción del último motivo (R5). La presencia de R5 al final de la secuencia de Msh43 es una característica común de todas las especies con excepción del gibón. Los últimos seis motivos de la secuencia de orangután están bien conservados en chimpancé (una transversión), gorila y humano (dos transiciones); por el contrario, en la organización del extremo 5' no se observa semejante grado de homología. Los primeros nueve motivos de la secuencia de Msh43 de humano y chimpancé son casi idénticos (una transición), mientras que el minisatélite de gorila presenta dos o tres cambios en este fragmento dependiendo del alelo. Por el contrario, en el extremo 3' es mayor la homología existente entre humano y gorila que con chimpancé.

El análisis de las distintas organizaciones de Msh43 muestra la existencia de abundantes duplicaciones en todos los alelos de las diferentes especies, lo que sugiere la implicación de procesos de *slippage* en la generación de polimorfismo en este locus.

En general, Msh43 presenta una organización asimétrica, en el sentido de que la distribución y el número de unidades repetitivas distintas entre las mitades 5' y 3' es diferente, hallándose la mayoría de las variantes de unidades repetitivas concentradas en la mitad 3' de la secuencia. Además, Msh43 es una secuencia heterogénea, puesto que hay bastantes unidades repetitivas distintas, algunas de éstas características de una especie.

El análisis del contenido de unidades repetitivas mostró que R1 (amarilla) es siempre la más abundante, seguida de R2 (roja) y R3 (verde) y que hay una inversión en la proporción R2/R3 en los MVM con respecto a gibón, Grandes Monos y humanos. (Figura 38).

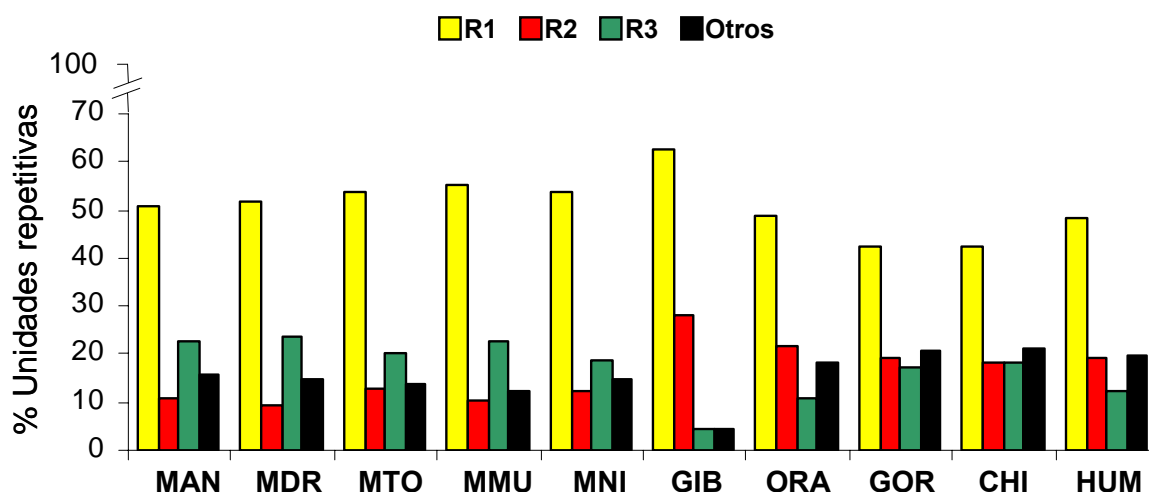


Figura 38. Histograma que muestra la abundancia relativa de las diferentes unidades repetitivas e MSH43. Los motivos R1, R2 y R3 están representados en amarillo, rojo y verde, siguiendo el código de color empleado en la figura 2. La barra negra representa el resto de las unidades repetitivas y que denominamos en el histograma como *otros*. Las abreviaturas de las diferentes especies empleadas son las mismas que las indicadas en la figura 37.

El análisis de la heterogeneidad (número de motivos diferentes con respecto al número total) indicó que el valor más alto correspondía a los Grandes Monos y humanos y el más bajo a gibón ya que sólo presenta seis unidades repetitivas distintas en su secuencia (Figura 39). Por otro lado, para establecer las relaciones entre los motivos se empleó el programa Network, mediante el algoritmo *Median Joining* y limitando los motivos incluidos en el análisis a aquellos presentes en más de una especie, evitando la aparición de ramas secundarias producidas por la existencia de motivos muy divergentes. El resultado de este estudio sitúa a R1 en la posición central

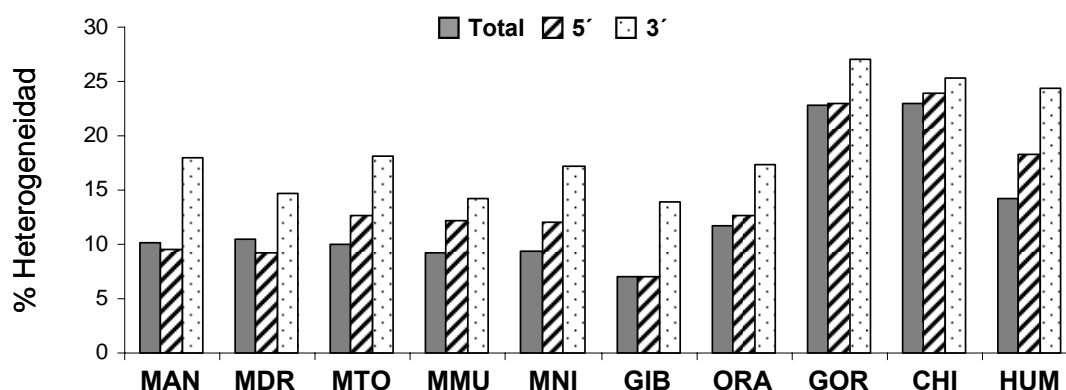


Figura 39. Histograma en el que se muestra la heterogeneidad de las unidades repetitivas en los alelos de Msh43 en los distintos géneros de primates. Este parámetro fue calculado como la proporción de los distintos motivos presentes en un alelo con respecto al número total (barra gris). La segunda y tercera barra indican la cantidad de heterogeneidad acumulada en las mitades 5' y 3' de Msh43, respectivamente.

de la red, dando lugar a R2 mediante una delección y a R3 a través de una transición (Figura 40). Este resultado, junto con el hecho de que R1 está ya presente en los primates ancestrales y que es la unidad repetitiva más abundante en todas las especies estudiadas, sugiere que este motivo constituye la secuencia inicial que posteriormente se repitió para generar el minisatélite.

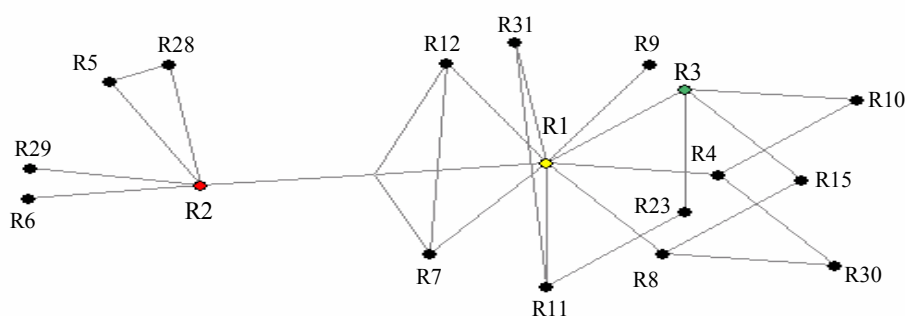


Figura 40. Relaciones entre las unidades repetitivas de Msh43. La nomenclatura de los motivos es la misma que la indicada en la figura 2. Solamente los motivos R1, R2 y R3 se representan en su correspondiente código de color.

Una diferencia destacable entre el locus Msh43 existente en Grandes Monos africanos y humanos en comparación con el de orangután, gibón y MVM, es una delección de ~ 1.6 Kb que abarca desde la posición 137 a la 1791 del locus Msh43 de orangután, incluyendo una secuencia *Alu* y los 35 primeros motivos del minisatélite

MsH43 existente en esta especie (Figura 41). La comparación entre las secuencias MsH43 de humano y orangután muestra la existencia de un alto grado de homología (sólo hay una mutación puntual) entre los 18 primeros motivos del MsH43 humano y el fragmento del minisatélite de orangután que se corresponde con las unidades repetitivas 36-53. Este resultado proporciona una prueba de que las 35 primeras unidades repetitivas del MsH43 de orangután estaban incluidas en la delección que tuvo lugar durante la evolución de los Grandes Monos.

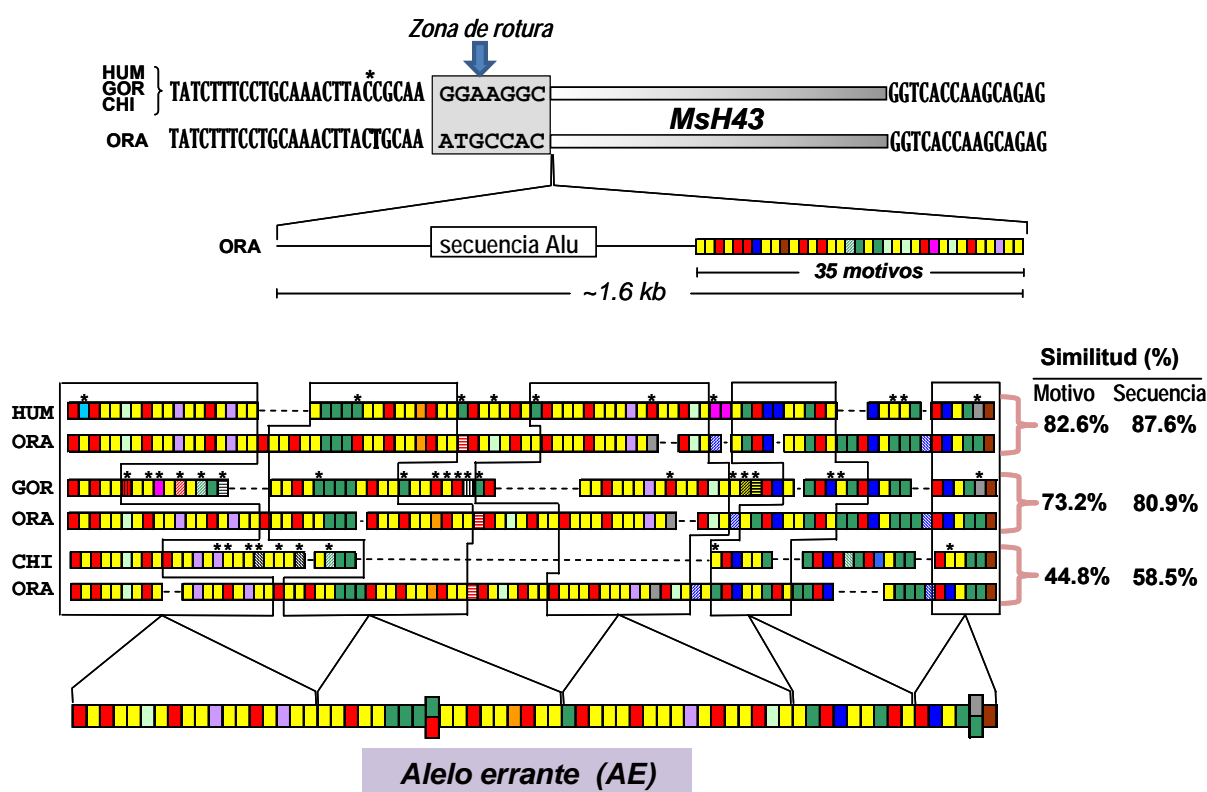


Figura 41. Representación esquemática de la comparativa de MsH43 y sus secuencias flanqueantes entre ORA y HUM, CHI y GOR. La zona de rotura (caja sombreada) corresponde a un fragmento de 7 pb en las que la secuencia de ORA es diferente a la que presentan HUM, CHI y GOR, y que se encuentra situada al final de la secuencia que flanquea el extremo 5' de MsH43. Se representa la estructura del fragmento de ORA que fue deleccionado en HUM, CHI y GOR. También se muestra el alineamiento del MsH43 de ORA (a partir del motivo 36) con las secuencias de HUM, CHI y GOR. Las líneas punteadas se han incluido para facilitar la comparación entre las secuencias; los asteriscos indican mutaciones en la unidad repetitiva. Los valores de similitud se han calculado comparando las secuencias nucleotídicas de MsH43, usando el alineamiento generado por el programa Bioedit (Secuencia) y la organización repetitiva (Motivo). Se ha construido la hipotética secuencia del AE, teniendo en cuenta las regiones de MsH43 donde la homología entre las cuatro especies es elevada (cajas).

Cabe destacar que el alineamiento de las secuencias que flanquean el extremo 5' de MsH43 de los Grandes Monos y humanos, pone de manifiesto la existencia de un fragmento de 7 pb que presenta una secuencia diferente en orangután con respecto a los demás. Este fragmento, que denominamos zona de rotura, está localizado justo antes del inicio del minisatélite en humano, chimpancé y gorila (Figura 41) y marca la posición de la rotura en el cromosoma ancestral. Las diferencias nucleotídicas en esta zona podrían reflejar el proceso de reparación que tiene lugar tras la rotura. Por otro lado, en contra de lo esperado a partir de las relaciones evolutivas consenso entre Grandes Monos y humanos, la comparación de la organización de MsH43 mostró que el mayor grado de homología lo presentan humano y orangután (Figura 41), lo cual no sigue las relaciones filogenéticas estándar. Para explicar esta inusual relación proponemos la existencia de una forma alélica conservada desde el antecesor común de Grandes Monos y humanos, que denominamos alelo errante y cuya posible secuencia se muestra en la Figura 41.

¿Cómo pudo desarrollarse la historia evolutiva de MsH43? El nacimiento de MsH43 se puede situar antes de la divergencia de MVM y después de la aparición de MNM (Figura 42). Esta hipótesis se basa en la existencia de algunas unidades repetitivas en Prosimios (R3-R28) y MNM (R1-R4-R5 y R8-R4-R5), aunque estas secuencias no presentan una organización repetitiva puesto que todos los motivos son diferentes. Sin embargo, su presencia en otros primates sugieren que esta secuencia podría representar una estructura preminisatélite, que sería el material genético inicial a partir del cual se construyó MsH43. Estos motivos pudieron haber sido generados por mutaciones al azar de una secuencia progenitora dando lugar al preminisatélite MsH43 (pre-MsH43). Es posible que la secuencia de dicho pre-MsH43 tuviese una organización muy similar a la secuencia del locus MsH43 de los MNM ya que las unidades repetitivas encontradas en estos (R1, R4, R5 y R8) se conservaron en todos los primates con la excepción del gibón que sólo presenta el motivo R1. Una característica interesante de la secuencia MsH43 de MNM y las secuencias que lo flanquean es que contienen varios fragmentos de DNA como palíndromos (CAGCTG, GAGCTC) y pseudopalíndromos (AGACT, AGGCT, etc) que pueden actuar como señales estructurales. La presencia de estas señales en el locus pre-MsH43 pudo inducir procesos de *slippage* durante la replicación que habrían tenido dos consecuencias: el crecimiento de la secuencia inicial, generando una organización en tándem, y la

aparición de nuevas unidades repetitivas como resultado de errores en la reparación de los heterodúplex que se generan durante el *slippage*.

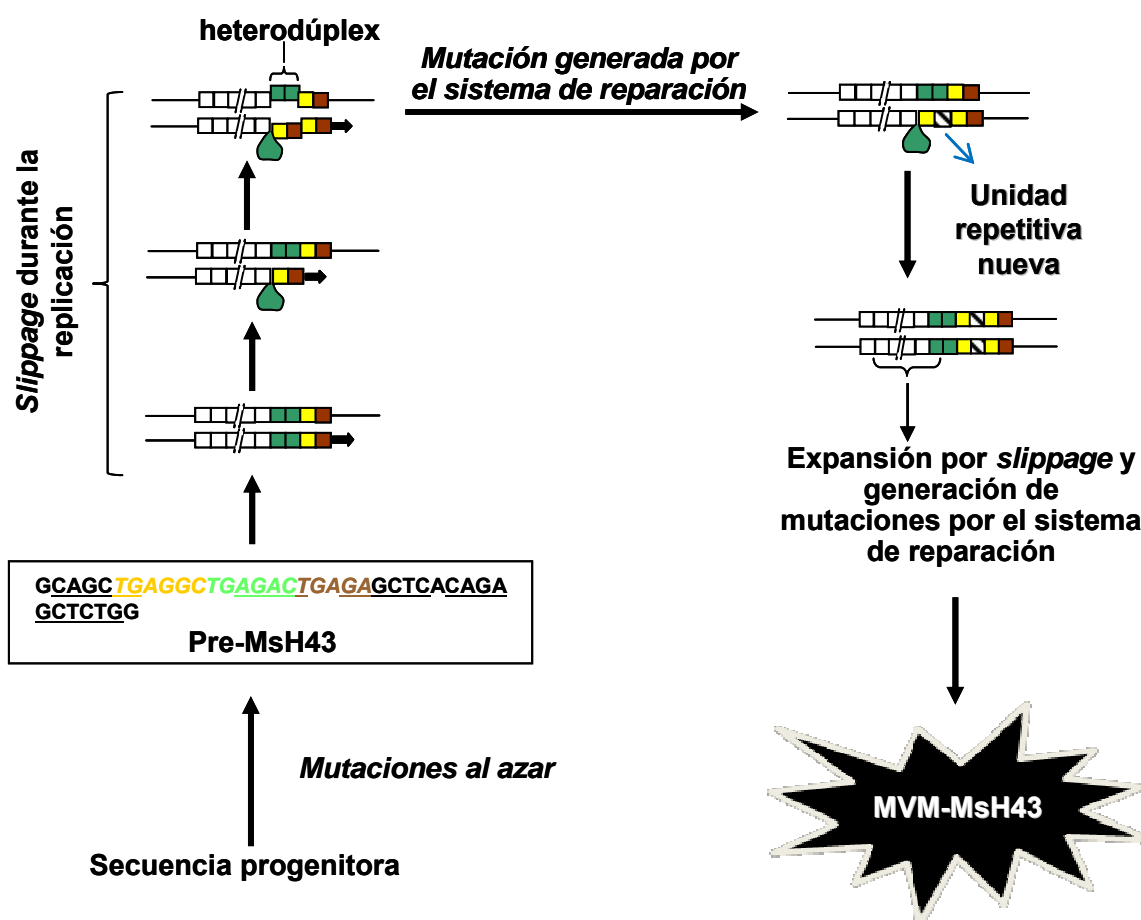


Figura 42. Esquema que representa el hipotético surgimiento de MSH43 en MVM. Éste tendría lugar a partir de una secuencia progenitora que por azar acumuló mutaciones que originaron el preminisatélite MSH43 (pre-MsH43). La presencia de estas secuencias provocó eventos de *slippage* que dieron lugar al crecimiento del pre-MsH43, generando una organización en tándem y la aparición de nuevas unidades repetitivas producidas por errores en el sistema de reparación. Los palíndromos y pseudopalíndromos se encuentran subrayados en la secuencia del pre-MsH43.

Apoyando esta hipótesis, ha sido descrito que secuencias que contienen duplicaciones en tándem son proclives a sufrir procesos de *slippage* (Buschiazzi & Gemmel, 2006; Rose & Falush, 1998). Además, existen diversos resultados que apoyan la implicación de las secuencias palindrómicas en el desencadenamiento de eventos de *slippage*. Así, han sido asociadas a la generación de cambios que pueden producir inestabilidad y reorganizaciones en genomas de una importante variedad de organismos eucariotas (Lewis & Coté, 2006). En este sentido, se ha descrito la participación de las secuencias

palindrómicas en varias clases de reorganizaciones genómicas ilegítimas y eventos de recombinación homóloga que dan lugar a la expansión de unidades repetitivas (Zhou et al., 2001), probablemente actuando de señales para la generación de cortes de doble cadena. También se han asociado las secuencias palindrómicas con la generación de inestabilidad mediada por *slippage* sobre la cadena retardada de la horquilla de replicación entre dos secuencias cortas repetidas directas (Leach, 1994). El mecanismo probablemente sea la yuxtaposición de las secuencias repetidas directas por emparejamiento de bases intracatenario. Una vez se han generado duplicaciones en tándem, estas secuencias cortas podrían sufrir eventos de *slippage* (Rose & Falush, 1998). En el caso de Msh43 los datos sugieren un modelo que consiste en la producción por azar de unidades palindrómicas iniciales en la secuencia progenitora, originando el pre-Msh43, que podría ser la señal para una combinación de eventos de *slippage* y/o entrecruzamientos desiguales, asociados con algunas mutaciones puntuales. Varios de estos eventos podrían aumentar el número de los motivos repetitivos iniciales a lo largo de millones de años, generando el nacimiento de Msh43 antes de la divergencia de los MVM. Existen dos datos que sugieren que la expansión de Msh43 provocada por el *slippage* durante la replicación pudo haber ocurrido hacia el extremo 5' en la secuencia Msh43 primitiva. Primero, en todas las especies analizadas, excepto gibbon, el último motivo es R5, y segundo, la heterogeneidad de Msh43 es siempre mayor en el extremo 3'. Esta polaridad respecto a la tasa de mutación ha sido detectada en otras secuencias repetitivas en tándem (Jeffreys et al., 1994; Murray et al., 1999; Ellegren, 2000). Una posible explicación es que la secuencia minisatélite sufra una tendencia a acumular procesos recombinogénicos en uno de sus extremos (Brohede & Ellegren, 1999). Por otro lado, un sesgo similar en eventos de *slippage* durante la replicación (Sia et al., 1997), también podría ser el mecanismo responsable de esta acumulación de motivos distintos en el extremo 3' de Msh43. La reparación de los heterodúplex generados por el *slippage* introduciría mutaciones en la secuencia de Msh43 dando lugar a nuevas unidades repetitivas tal como ha sido descrito para otras secuencias (Buschiazzo & Gemmel, 2006; Ellegren, 2004). Este modelo de expansión de Msh43 y generación de nuevos motivos por el sistema de reparación es compatible con el hecho de que los bucles desplazados de la cadena de nueva síntesis, resultantes del *slippage* de la DNA polimerasa, se reparan de forma distinta que los bucles de la cadena molde (Sia et al., 1997).

Es conocido que los genomas de humanos y de otros primates todavía recuerdan a su antecesor común, sin embargo el genoma del gibón parece haber sufrido una reorganización masiva (Roberto et al., 2007). Por esto, no resulta sorprendente que la estructura de MsH43 en gibón sea peculiar, presentando una arquitectura similar a la de un micro/minisatélite compuesto por la combinación de dos unidades repetitivas. Basándonos en el alto grado de homología que presentan los extremos 5' de MVM y gibón (Figura 43), y en la conservación de las secuencias que flanquean los extremos 3', sugerimos que en el linaje que conduce a gibón tuvo lugar una delección en el locus MsH43 que implicó a una región interna del minisatélite, dejando una zona, que denominamos MsH43i, homóloga a la región 5' de MVM rica en R1 y R2. Esa secuencia de MsH43 que permaneció, pudo haber sufrido varios procesos de *slippage* homogenizador (no se generaron motivos nuevos) que provocaron su expansión. Puesto que la tasa de mutación de un microsatélite aumenta con el número de unidades repetitivas (Buschiazzi & Gemmel, 2006; Jurka & Gentles, 2006; Webster et al., 2002) es razonable esperar que durante la expansión de MsH43 hacia el extremo 3', apareciesen nuevas motivos como consecuencia de errores del sistema de reparación de los heterodúplex generados por *slippage*.

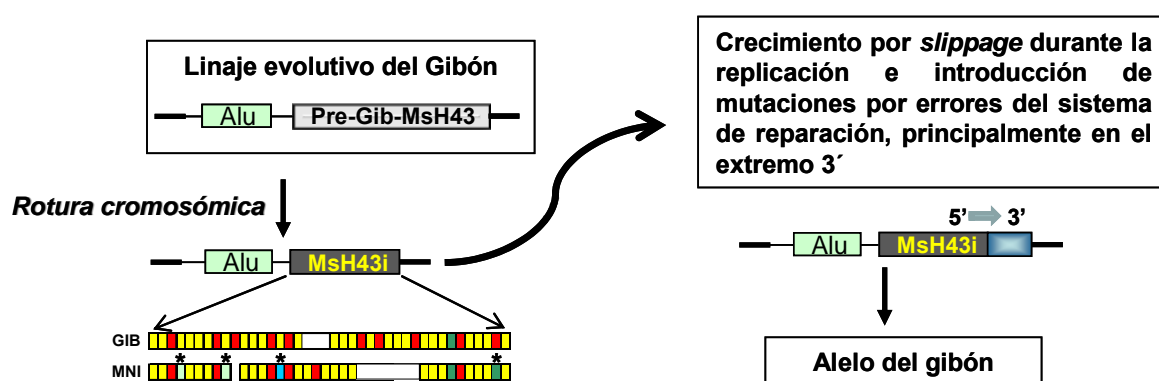


Figura 43. Génesis del alelo de gibón. A partir de los datos experimentales disponibles asumimos la existencia de un intermediario evolutivo, el Pre-Gib-MsH43, que contiene una secuencia *Alu* en la zona 5' adyacente de MsH43 y tendría una organización similar al MsH43 de MVM. Cabe destacar el alto grado de homología en la organización de unidades repetitivas entre los extremos 5' del actual MsH43 de gibón y MNI. A partir de esta comparación, deducimos que tuvo lugar una rotura en el cromosoma que portaba MsH43. Posteriormente, esta secuencia sufrió varios eventos de *slippage* que produjeron el crecimiento de MsH43 hacia el extremo 3'. Durante este proceso se generaron nuevos motivos como consecuencia de errores del sistema de reparación.

Tras la divergencia entre orangután y Grandes Monos africanos y humano, el locus MsH43 sufrió una gran modificación: una rotura en el cromosoma que alberga esta región que produjo la delección de una zona conteniendo la secuencia flanqueante a 5' de MsH43, que incluye una *Alu*, y las 35 primeras unidades repetitivas del minisatélite. En este sentido, se ha descrito que las secuencias *Alu* pueden verse implicadas en la generación de delecciones y/o duplicaciones de fragmentos de DNA en el genoma de células meióticas, mediante procesos de recombinación homóloga desigual (Kolomietz et al., 2002). Así, se ha visto que varias delecciones observadas en algunos tipos de reorganizaciones cromosómicas asociados con leucemia ocurren con más probabilidad cuando las regiones que sufren reorganizaciones contienen una gran cantidad de secuencias *Alu* ((Kolomietz et al., 2002). Además, el análisis de los intervalos de rotura observada en las microdelecciones que se detectan en el síndrome de la neurofibromatosis tipo1 sugiere que secuencias *Alu* podrían estar mediando directamente en esta delección (Gervasini et al., 2005).

La comparación de las secuencias de MsH43 entre Grandes Monos y humanos mostró que el mayor grado de homología en la organización de dicho locus lo presentan orangután y humano. Este resultado es muy sorprendente puesto que contradice la filogenia estándar de primates y cambia el paradigma evolutivo de que durante el curso de la evolución genómica se producen continuas alteraciones. Puesto que orangután y el linaje que da lugar a los Grandes Monos africanos y humanos divergieron hace aproximadamente 12-15 millones de años, sugerimos la existencia de un alelo, que denominamos alelo errante (AE), que se mantuvo con pequeñas variaciones a lo largo del período evolutivo que se extiende entre orangután y humano (Figura 44). La hipótesis del AE implica la existencia de un alelo que se mantuvo en algunos de los individuos que fueron intermediarios en la filogenia de primates. En cada rama evolutiva, el AE sufrió pequeños cambios hasta alcanzar la actual versión de MsH43 en las diferentes especies. Así, en el caso del gorila el AE evolucionó mediante duplicaciones y mutaciones puntuales, mientras que en el linaje del chimpancé probablemente sufrió una delección (los alelos de chimpancé son más pequeños que los de las demás especies) y también algunas duplicaciones y mutaciones puntuales. En la especie humana, la comparación entre los alelos muestra que las diferencias entre ellos son de tamaño, debidas a duplicaciones y delecciones generadas posiblemente por *slippage* durante la replicación.

El análisis de la organización de los últimos tres motivos en las secuencias de los Grandes Monos africanos y humanos pone de manifiesto un dato interesante: mientras que en orangután y chimpancé el orden de las unidades repetitivas es R3-R3-R5, en gorila y humano es R3-R8-R5. La existencia de esta mutación testigo sugiere que pudo haber al menos dos variantes del AE. De acuerdo con esto, una de las variantes del AE se mantuvo en el linaje que condujo a gorila y humano y la otra variante estuvo implicada en el linaje evolutivo que dio lugar a chimpancé.

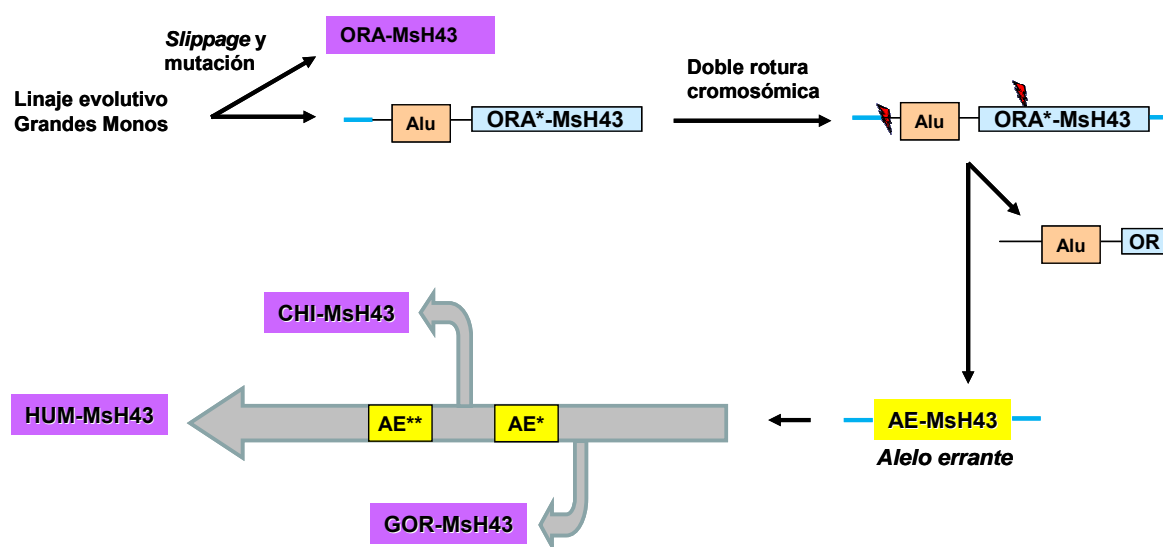


Figura 44. Generación de MSH43 en Grandes Monos y humanos: la hipótesis del alelo errante. En el locus MSH43 que presentaba una organización similar al actual MSH43 de ORA (marcado con un asterisco), tuvo lugar una doble rotura. Este corte produjo la pérdida de un fragmento de la secuencia que flanquea a MSH43 por el extremo 5', incluyendo una secuencia *Alu* y los 35 primeros motivos de MSH43 de ORA. Después de este suceso, se generó una nueva forma alélica, el AE, que persistió hasta los humanos actuales. Este AE ha sufrido algunas mutaciones durante el curso de la evolución generando variantes del AE (señaladas con asterisco) pero su estructura permaneció prácticamente inalterada. En el caso de GOR, el AE probablemente evolucionó a través de duplicaciones y mutaciones puntuales, mientras que el CHI posiblemente sufrió una delección, que fue el cambio más importante. En el antecesor de las especies humanas, el AE llegó con pocas modificaciones de acuerdo con la homología observada entre el MSH43 de HUM y ORA.

Varios son los factores (entrecruzamiento desigual, delecciones mediadas por secuencias *Alu*, mutaciones al azar, roturas en cromosomas, *slippage* durante la replicación y errores en el sistema de reparación) que parecen haber influido en el

destino de Msh43 a través de la evolución, y la existencia de características específicas de un taxón es consecuencia de combinaciones únicas de estos.

Finalmente, de acuerdo con nuestro modelo, la estabilidad de Msh43 en primates podría ser la consecuencia de eventos de *slippage* durante la replicación que tuviesen dos efectos mutagénicos opuestos: el crecimiento de Msh43 y la generación de mutaciones en la organización repetitiva que tenderían a romper la estructura de minisatélite. La producción de roturas de doble cadena y/o eventos de entrecruzamiento desigual producirían los principales cambios de Msh43 durante la especiación.

CONCLUSIONES

El análisis de los resultados obtenidos en la presente Tesis Doctoral permite extraer las siguientes conclusiones:

1. El estudio del polimorfismo del minisatélite MSH43 nos permitió identificar ocho alelos. La amplificación de MSH43 de individuos heterocigotos genera estructuras heterodúplex que les confiere un patrón electroforético específico que hemos empleado en la identificación de nuevos genotipos.
2. La mayoría de los alelos de MSH43 que hemos aislado presentan la capacidad de formar estructuras G-cuadrúplex que son estabilizadas por potasio, y que pueden interferir en la síntesis de DNA durante la PCR.
3. El incremento de la fuerza iónica en los ensayos de recombinación homóloga intramolecular *in vitro* genera una disminución de la actividad nucleásica y produce una mayor fidelidad en los eventos recombinatorios puesto que aumenta la proporción de moléculas recombinantes que mantienen la secuencia original.
4. A partir de la secuencia del locus MSH43 de diferentes especies de primates hemos reconstruido la historia evolutiva de este minisatélite. Así, encontramos que esta secuencia se originó durante la génesis de los Monos del Viejo Mundo a partir de un preminisatélite que apareció en los Monos del Nuevo Mundo. La homología encontrada entre humano y orangután sugiere la existencia de un alelo, que denominamos alelo errante, que ha permanecido prácticamente inalterado durante el período evolutivo comprendido entre orangután y humano.

BIBLIOGRAFÍA

- Ali, S. & Wallace, R.B.** (1988) Intrinsic polymorphism of variable number tandem repeat loci in the human genome. *Nucleic Acids Res.* **16**, 8487-8496.
- Almeida, P. & Penha, C.** (2004) Long perfect dinucleotide repeats are typical of vertebrates show motif preferences and size convergence. *Mol. Biol. Evol.* **21**, 1226-1233.
- Amos, W., Sawcer, S.J., Feakes, R.W. & Rubinsztein, D.C.** (1996) Microsatellites show mutational bias and heterozygote instability. *Nat. Genet.* **13**, 390-391.
- Andersen T.H. & Nilsson-Tillgren T.** (1997) A fungal minisatellite. *Nature* **386**, 771.
- Andreassen, R., Lundsted, J., & Olaisen, B.** (2002) Mutation at minisatellite locus DYF155S1: Allele length mutation rate is affected by age of progenitor. *Electrophoresis* **23**, 2377-2383.
- Anglani, F., Picci, L., Camporese, C. & Zacchello, F.** (1990) Heterodúplex formation in polymerase chain reaction. *Am. J. Hum. Genet.* **47**, 169-170.
- Armour, J.A.L., Harris, P.C. & Jeffreys, A.J.** (1993) Allelic diversity at minisatellite MS205 (D16S309): evidence for polarized variability. *Hum. Mol. Genet.* **2**, 1137-1145.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D.D., Seidman, J. G., Smith, J. A. & Struhl, K.** (1987) *Current Protocols in Molecular Biology*. Wiley Interscience, NY.
- Beckmann, J.S. & Weber, J.L.** (1992) Survey of human and rat microsatellites. *Genomics* **12**, 627-631.
- Bennett, S.T., Lucassen, A.M., Gough, S.C., Powell, E.E., Undlien, D.E., Pritchard, L.E. et al.** (1995) Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat. Genet.* **9**, 284-292.
- Berg, I., Neumann, R., Cederberg, H., Rannug, U. & Jeffreys, A.J.** (2003) Two modes of germline instability at human minisatellite MS1 (locus D1S7): complex rearrangements and paradoxical hyperdeletion. *Am. J. Hum. Genet.* **72**, 1436-1447.
- Bhattacharyya, A & Lilley, D.M .** (1989) The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). *Nucleic Acids Res.* **17**, 6821-6840.
- Blanco, M.G., Boán, F, Barros, P., Castaño J.G. & Gómez-Márquez, J.** (2005) Generation of DNA double-strand breaks by two independent enzymatic activities in nuclear extracts. *J. Mol. Biol.* **351**, 95-1006.
- Blanco, M.G., Boán, F. & Gómez-Márquez, J.** (2004) A paradox in the *in vitro* end-joining assays. *J. Biol. Chem.* **279**, 26797-26801.

- Boán, F., Blanco M.G., Barros, P. & Gómez-Márquez, J.** (2006) DNA end-joining driven by microhomologies catalyzed by nuclear extracts. *Biol. Chem.* **387**, 263-267.
- Boán, F., Blanco, M.G., Quinteiro, J., Mouriño, S., Gómez-Márquez, J.** (2004) Birth and evolutionary history of a human minisatellite. *Mol. Biol. Evol.* **21**, 228-235.
- Boán, F., González, A.I., Rodríguez, J.M. & Gómez-Márquez, J.** (1997) Molecular characterization of a new human minisatellite that is able to form single-stranded loops *in vitro* and is recognized by nuclear proteins. *FEBS Lett.* **418**, 251-257.
- Boán, F., Rodríguez, J.M. & Gómez-Márquez, J.** (1998) A non-hypervariable human minisatellite strongly stimulates *in vitro* intramolecular homologous recombination. *J. Mol. Biol.* **278**, 499-505.
- Boán, F., Rodríguez, J. M., Mouriño, S., Blanco, M.G., Viñas, A., Sánchez, L. & Gómez-Márquez, J.** (2002) Recombination analysis of the human minisatellite Msh42 suggests the existence of two distinct pathways for initiation and resolution of recombination at Msh42 in rat testes nuclear extracts. *Biochemistry* **41**, 2166-2176.
- Boguski, M.S., Birkenmeier, E.H., Elshourbagy, N.A., Taylor, M.T. & Gordon, J.I.** (1986) Evolution of the apolipoproteins. *J. Biol. Chem.* **261**, 6398-6407.
- Bois, P. & Jeffreys, A. J.** (1999) Minisatellite instability and germline mutation. *Cell. Mol. Life. Sci.* **55**, 1636-1648.
- Borstnik, B. & Pumpernik, D.** (2002) Tandem repeats in protein coding regions of primate genes. *Genome Res.* **12**, 909-915.
- Brinkmann, B., Klintschar, M., Neuhuber, F., Huhne, J. & Rolf, B.** (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.* **62**, 1408-1415.
- Britten, R.J. & Kohne, D.E.** (1968) Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* **161**, 529-540.
- Brock, G.J., Anderson, N.H. & Monckton, D.G.** (1999) Cis-acting modifiers of expanded CAG/CTG triplet repeat expandability: associations with flanking GC content and proximity to CpG islands. *Hum. Mol. Genet.* **8**, 1061-1067.
- Brohede, J. & Ellegren, H.** (1999) Microsatellite evolution: polarity of substitutions within repeats and neutrality of flanking sequences. *Proc Biol Sci* **266**, 825-833.
- Brown, T.A.** (1999) Genomes. *BIOS Scientific Publishers*. Oxford.
- Buard, J. & Vergnaud G.** (1994) Complex recombination events at the hypermutable minisatellite CEB1 (D2S90). *EMBO J.* **13**, 3203-3210
- Buard, J., Bourdet, A., Yardley, J., Dubrova, Y. & Jeffreys, A. J.** (1998) Influences of array size and homogeneity on minisatellite mutation. *EMBO J.* **17**, 3495-3502.

- Buard, J., Shone, A.C. & Jeffreys, A.J.** (2000) meiotic recombination and flanking marker exchange at the highly unstable human minisatellite CEB1 (DS290). *Am. J. Hum. Genet.* **67**, 333-344.
- Buschiazzo, E. & Gemmel, N.** (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays* **28**, 1040-1050.
- Cederberg, H. & Rannug, U.** (2006) Mechanisms of human minisatellite mutation in yeast. *Mutat. Res.* **598**, 132-143.
- Chaillet, J.R., Bader, D.S. & Jeder, P.** (1995) Regulation of genomic imprinting by gametic and embryonic processes. *Genes Dev.* **9**, 1177-1187.
- Chakraborty, R., Kimmel M., David N.S., Leslea J.D. & Deka, R.** (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc. Natl. Acad. Sci. USA* **94**, 1041-1046.
- Chandley, A.C. & Mitchell, A.R.** (1988) Hypervariable minisatellite regions are sites for crossing-over at meiosis in man. *Cytogenet. Cell Genet.* **48**, 152-155.
- Charlesworth, B., Sniegowski, P. & Stephan, W.** (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**, 215-220.
- Chen, F.M.** (1992) Sr^{2+} facilitates intermolecular G-quadruplex formation of telomeric sequences. *Biochemistry* **31**, 3769-3776.
- Collick, A. & Jeffreys, A.J.** (1990) Detection of a novel minisatellite-specific DNA binding protein. *Nucleic Acids Res.* **18**, 625-629.
- Crnugelj, M., Hud, N.V. & Plavec, J.** (2002) The solution structure of d(G(4)T(4)G(3))(2): a bimolecular G-quadruplex with a novel fold. *J. Mol. Biol.* **320**, 911-924.
- Dallas, J.F.** (1988) Detection of DNA "fingerprints" of cultivated rice by hybridization with a human minisatellite DNA probe. *Proc. Natl. Acad. Sci. USA* **85**, 6831-6835.
- Dieringer, D. & Schlötterer, C.** (2003) Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. *Genome Res.* **13**, 2242-2251.
- Dignam, J.D., Lebovitz, R.M. & Roeder, R.G.** (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475-1489.
- Ellegren, H.** (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. *Nat. Genet.* **24**, 400-402.
- Ellegren, H.** (2004) Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* **5**, 435-445.

- Etzioni, S., Yafe, A., Khateb, S., Weisman-Shomer, P. et al.** (2005) Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes. *J. Biol. Chem.* **280**, 26805-26812.
- Fang, G. & Cech, T.R.** (1993) The beta subunit of *Oxytricha* telomere-binding protein promotes G-quarter formation by telomeric DNA. *Cell* **74**, 875-885.
- Fry, M.** (2007) Tetraplex DNA and its interacting proteins. *Front. Biosci.* **12**, 4336-4351.
- Fukuda, H., Takashi, S., Nagao, M. & Nakagama, H.** (2001) Detection and isolation of minisatellite Pc-1 binding proteins. *Biochim. Biophys. Acta* **1528**, 152-158.
- Gellert, M., Lipsett, M.N. & Davies, D.R.** (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA.* **48**, 2013-2018.
- Gervasini, C., Venturin, M., Orzan, F., Friso, A., Clemente, M., Tenconi, R., Larizza, L. & Riva, P.** (2005) Uncommon Alu-mediated NF-1 microdeletion with a breakpoint inside the NF-1 gene. *Genomics* **85**, 273-279.
- Gill, P., Jeffreys, A.J. & Werret, D.J.** (1985) Forensic applications of DNA "fingerprints". *Nature* **318**, 577-579.
- Gómez-Márquez, J., Puga, A. & Notkins, A.L.** (1985) Regions of the terminal repetitions of the *Herpes simplex* virus type I genome. *J. Biol. Chem.* **260**, 3490-3495.
- González, A.I.** (1993) Homología entre el DNA humano y el fragmento SmaIA del genoma del virus Herpes simplex tipo 1: nuevos minisatélites en el genoma humano. *Tesis Doctoral*. Departamento de Bioquímica e Biología Molecular, Facultad de Biología, Universidade de Santiago de Compostela.
- Gusmao, L., Sánchez-Diz, P., Calafell, F., Martín, P., Alonso, C.A. et al.** (2005) Mutation rates at Y chromosome specific microsatellites. *Hum. Mutat.* **26**, 520-528.
- Haddad, L.A., Fuzikama, A.K. & Pena, S.D.** (1997) Simultaneous detection of size and sequence polymorphisms in the transcribed trinucleotide repeat D2S196E (EST00493). *Hum. Genet.* **99**, 796-800.
- Hall, T.A.** (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**, 95-98.
- Han, H., Hurley, L.H. & Salazar, M.** (1999) A DNA polymerase stop assay for G-quadruplex-interactive compounds. *Nucleic Acid. Res.* **27**, 537-542.
- Harr, B., Werss, S., David, J.R., Brem, G. & Schlötterer, C.** (1998) A microsatellite-based multilocus phylogeny of the *Drosophila melanogaster* species complex. *Curr. Biol.* **8**, 1183-1186.

- Harrington, C., Lan, Y. & Akman, S.A.** (1997) The identification and characterization of a G4-DNA resolvase activity. *J. Biol. Chem.* **272**, 24631-24636.
- He, Q., Cederberg, H. & Rannug, U.** (2002) The influence of sequence divergence between alleles of the human MS205 minisatellite incorporated into the yeast genome on length-mutation rates and lethal recombination events during meiosis. *J. Mol. Biol.* **319**, 315-327.
- Heller, M., Henderson, A. & Kieff, E.** (1982) Repeat array in Epstein-Barr virus DNA is related to cell DNA sequences interspersed on human chromosomes. *Proc. Natl. Acad. Sci. USA* **79**, 5916-5920.
- Henke J. & Henke L.** (1995) Recent observations in human DNA-minisatellite mutations. *Int. J. Legal Med.* **107**, 204-208.
- Hite, J.M., Eckert, A. & Cheng, K.C.** (1996) Factors affecting fidelity of DNA synthesis during PCR amplification of d(C-A)_n. d(G-T)_n microsatellite repeats. *Nucleic Acids Res.* **24**, 2429-2434.
- Hoffman, E.K., Trusko, S.P., Murphy, M. & George, D.L.** (1990) An S1 nuclease-sensitive homopurine/homopyrimidine domain in the c-kis-ras promoter interacts with a nuclear factor. *Proc. Natl. Acad. Sci. USA* **87**, 2705-2709.
- Huang, Q.Y., Xu, F.H., Shen, H., Deng, H.Y., Liu, Y.J. et al.** (2002) Mutation patterns at dinucleotide microsatellite loci in humans. *Am. J. Hum. Genet.* **70**, 625-634.
- Hud, N.V., Schultze, P. & Feigon, J.** (1998) Ammonium ion as a NMR probe for monovalent cation coordination sites of DNA quadruplexes. *J. Am. Chem. Soc.* **120**, 6403-6404.
- Hunt, P.A. & Hassold, T.J.** (2002) Sex matters in meiosis. *Science* **296**, 2181-2183.
- Inoue, H., Nojima, H. & Okayama, H.** (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-28.
- International Human Genome Sequencing Consortium** (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.
- International Human Genome Sequencing Consortium** (2004) Finishing the euchromatic sequencing of the human genome. *Nature* **431**, 931-945.
- Jarne, P. & Lagoda, P.J.L.** (1996) Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11**, 424-429.
- Jauert, P.A. & Kirkpatrick, D.T.** (2005) Length and sequence heterozygosity differentially affect HRAS1 minisatellite stability during meiosis in yeast. *Genetics* **170**, 601-612.

- Jeffreys, A.J., Brookfield, J.F.Y. & Semeonoff, R.** (1985c) Positive identification of an immigration test-case using human DNA fingerprints. *Nature* **317**, 818-819.
- Jeffreys, A.J. & Morton, D.B.** (1987) DNA fingerprints of dogs and cats. *Anim. Genet.* **18**, 1-15.
- Jeffreys, A.J., Neil, D.L. & Neumann, R.** (1998) Repeat instability at human minisatellites arising from meiotic recombination. *EMBO J.* **17**, 4147-4157.
- Jeffreys, A.J., Neumann, R. & Wilson, V.** (1990) Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell* **60**, 473-485.
- Jeffreys, A.J., Tamaki, K., MacLeod, A., Monckton, D.G., Neil, D.L. & Armour, J.A.L.** (1994) Complex gene conversion events in germline mutation at human minisatellites. *Nat. Genet.* **6**, 136-145.
- Jeffreys, A.J., Wilson, V. & Thein, S.L.** (1985a) Hypervariable minisatellite regions in human DNA. *Nature* **314**, 67-73.
- Jeffreys, A.J., Wilson, V. & Thein, S.L.** (1985b) Individual-specific "fingerprints" of human DNA. *Nature* **316**, 76-79.
- Jobling, M.A. & Gill P.** (2004) Encoded evidence: DNA in forensic analysis. *Nat. Rev. Genet.* **5**, 739-751.
- Jurka, J. & Gentles, A.J.** (2006) Origin and diversification of minisatellites derived from human Alu sequences. *Gene* **365**, 21-26.
- Kolomietz, E., Meyn, M.S., Pandita, A. & Squire, J.A.** (2002) The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. *Genes Chromosomes Cancer* **35**, 97-112.
- Kreis, S. & Whistler, T.** (1997) Rapid identification of measles virus strains by the heteroduplex mobility assay. *Virus Res.* **47**, 197-203.
- Lai, Y. & Sun, F.** (2003) The relationship between microsatellite slippage mutation rate and the number of repeat units. *Mol. Biol. Evol.* **20**, 2123–2131.
- Lalioti, M.D., Scott, H.S., Buresi, C., Rossier, C., Bottani, A., Morris, M.A. et al.** (1997) Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. *Nature* **386**, 847-850.
- Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. & Maniatis, T.** (1978) The isolation and characterization of linked delta- and beta-globin genes from a cloned library of human DNA. *Cell* **15**, 1157-1174.
- Leach, D.R.** (1994) Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* **16**, 839-900.

- Leung, W.K., Kim, J.J., Kim, J.G., Graham, D.Y. & Sepulveda, A.R.** (2000) Microsatellite instability in gastric intestinal metaplasia in patients with and without gastric cancer. *Am. J. Pathol.* **156**, 537-543.
- Levinson, G. & Gutman, G.A.** (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**, 203-221.
- Lewin, B.** (2008) Genes IX. *Jones and Bartlett Publishers*. Massachusetts.
- Lewis, S.M. & Cote, A.G.** (2006) Palindromes and genomic stress fractures: bracing and repairing damage. *DNA Repair* **5**, 1146-1160.
- Lichter, J.B., Barr, C.L., Kennedy, J.L., Van Tol, H.H.M., Kidd, K.K. & Livak, K.J.** (1993) A hypervariable segment in the human dopamine receptor D4 (D4DR) gene. *Hum. Mol. Genet.* **2**, 767-773.
- Liu, L., Dybvig, K., Panangala, V.S. van Santen V.L. & French, C.T.** (2000) GAA trinucleotiderepeat region regulates M9/pMGA gene expression in *Mycoplasma gallisepticum*. *Infect. Immun.* **68**, 871-876.
- Liu, Y., Stasiak, A.Z., Masson, J., McIlwraith, M.J., Stasiak, A. & West, S.C.** (2004) Conformational changes modulate the activity of human RAD51 protein. *J. Mol. Biol.* **337**, 817-827.
- Lopes, J., Ribeyre, C. & Nicolas, A.** (2006) Complex minisatellite rearrangements generated in the total or partial absence of Rad27/hFEN1 activity occur in a single generation and are Rad51 and Rad52 dependent. *Mol. Cell Biol.* **26**, 6675-6689.
- Maizels, N.** (2006) Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nature* **13**, 1055-1059.
- Majewski, J. & Ott, J.** (2000) GT repeats are associated with recombination on human chromosome 22. *Genome Res.* **10**, 1108-1114.
- Meloni, R., Albanese, V., Ravassard, P., Treilhou, F. & Mallet, J.** (1998) A tetranucleotidic polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, act as a transcription regulatory element *in vitro*. *Hum. Mol. Genet.* **7**, 423-428.
- Messier, W., Li, S.H. & Stewart, C.B.** (1996) The birth of microsatellites. *Nature* **381**, 483.
- Moncada, X., Pelsy, F., Merdinoglu, D. & Hinrichsen, P.** (2006) Genetic diversity and geographical dispersal in grapevine clones revealed by microsatellite markers. *Genome* **49**, 1459-1472.
- Monckton, D.G. Neumann, R., Guram, T., Fretwell, N., Tamaki, K., Macleod, A. & Jeffreys A.J.** (1994) Minisatellite mutation rate variation associated with a flanking DNA sequence polymorphism. *Nat. Genet.* **8**, 162-170

- Mouhamadou, B., Ferandon, C., Chazoule, S. & Barroso, G.** (2007) Unusual accumulation of polymorphic microsatellite loci in a specific region of the mitochondrial genome of two mushroom-forming *Agrocybe* species. *FEMS Microbiol. Lett.* **272**, 276-281.
- Mrazek, J., Guo, X & Shah, A.** (2007) Simple sequence repeats in prokaryotic genomes. *Proc. Natl. Acad. Sci. USA* **104**, 8472-8477.
- Murray, J., Buard, J., Neil, D.L., Yeramian, E., Tamaki, K., Hollies, C. & Jeffreys, A.J.** (1999) Comparative sequence analysis of human minisatellites showing meiotic repeat instability. *Genome Res.* **9**, 130-136.
- Nakamura, Y., Leppert, M., O'Connel, P., Wolff, R., Holm, R. et al.** (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**, 1616-1622.
- Neidle, S. & Read, M.A.** (2000) G-quadruplexes as therapeutic targets. *Biopolymers* **56**, 195-208.
- Neumann, B., Kubicka, P. & Bartel, D. P.** (1995) Characteristics of imprinted genes. *Nat. Genet.* **9**, 12-13.
- Orgel, L.E. & Crick, F.H.C.** (1980) Selfish DNA: the ultimate parasite. *Nature* **284**, 604-607.
- Pääbo, S.** (2003) The mosaic that is our genome. *Nature* **421**, 409-412.
- Pâques, F. & Haber, J.E.** (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **63**, 349-404.
- Peden, K., Mounts, P. & Hayward G.S.** (1982) Homology between mammalian cell DNA sequences and human herpesvirus genomes detected by a hybridization procedure with high-complexity probe. *Cell* **31**, 71-80.
- Perez, M., Cruz, F. & Presa, P.** (2005) Distribution properties of polymononucleotide repeats in molluscan genomes. *J. Hered.* **96**, 40-51.
- Puga, A., Cantin, E.M. & Notkins, A.L.** (1982) Homology between murine and human cellular DNA sequences and terminal repetition of the S component of herpes simplex virus type 1 DNA. *Cell* **31**, 81-87.
- Pupko, T. & Graur, D.** (1999) Evolution of microsatellites in the yeast *Saccharomyces cerevisiae*: role of length and number of repeated units. *J. Mol. Evol.* **48**, 313-316.
- Richard, G.F. & Paques, F.** (2000) Mini- and microsatellite expansions: the Recombination connection. *EMBO Rep.* **1**, 122-126.

Roberto, R., Capozzi, O., Wilson, R.K., Mardis, E.R., Lomiento, M., Tuzun, E. et al. (2007) Molecular refinement of gibbon genome rearrangements. *Genome Res.* **17**, 249-257.

Rodríguez, J.M. (2002) Estudio *in vitro* del minisatélite MsH42: Análisis poblacional y localización en el genoma humano, comparación de la capacidad recombinogénica de los alelos e interacción con proteínas nucleares. *Tesis Doctoral*. Departamento de Bioquímica e Biología Molecular, Facultad de Biología, Universidad de Santiago de Compostela.

Rodríguez, S. Visedo, G. & Zapata, C. (2001) Detection of errors in dinucleotide repeat typing by nondenaturing electrophoresis. *Electrophoresis* **22**, 2656-2664.

Rose, O. & Falush, D. (1998) A threshold size for microsatellite expansion. *Mol. Biol. Evol.* **15**, 613-615.

Royle, N.J., Clarkson, R.E., Wong, Z. & Jeffreys, A.J. (1988) Clustering of hypervariable minisatellites in the proterminal regions of the human autosomes. *Genomics* **3**, 352-360.

Saha, T. & Usdin, K. (2001) Tetraplex formation by the progressive myoclonus epilepsy type-1 repeat: implications for instability in the repeat expansion diseases. *FEBS Lett.* **491**, 184-187.

Sakano, H., Kurosawa, Y., Weigert, M. & Tonegawa S. (1981) Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. *Nature* **290**, 562-565.

Sambrook, J. & Russell D.W. (2001) Molecular cloning. A laboratory manual. *Cold Spring Harbor Laboratory Press*. New York.

Sen, D. & Gilbert, W. (1990) A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **344**, 410-414.

Shim, K. Schmutte, C., Yoder, K. & Fishel, R. Defining the salt effect on human RAD51 activities. *DNA repair* **5**, 718-730.

Shinde, D., Lai, Y., Sun, F. & Arnheim, N. (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)_n and (A/T)_n microsatellites. *Nucleic Acids Res.* **31**, 974-980.

Shukla, A.K. & Roy, K.B. (2006) RecA-independent homologous recombination induced by a putative fold-back tetraplex DNA. *Biol. Chem.* **387**, 251-256.

Sia, E.A., Kokoska, R., Dominska, M., Greenwell, P. & Petes, T.D. (1997) Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Mol. Cell Biol.* **17**, 2851-2858.

Sibly, R.M., Whittaker, J.C. & Talbot, M. (2001) A maximum-likelihood approach to fitting equilibrium models of microsatellite evolution. *Mol. Biol. Evol.* **18**, 413-417.

- Siddiqui-Jain, A., grand, C.L. Bearss, D.J. & Hurley, L.H.** (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA* **99**, 11593-11598.
- Simonsson, T.** (2001) G-quadruplex DNA structures--variations on a theme. *Biol. Chem.* **382**, 621-628.
- Smirnov, I. & Shafer, R.H.** (2000) Lead is unusually effective in sequence-specific holding of DNA. *J. Mol. Biol.* **296**, 1-5.
- Smith, G.P.** (1976) Evolution of repeated DNA sequences by unequal crossover. *Science* **191**, 528-535.
- Spink, C.F., Keen, L.J., Middleton, P.G. & Bidwell, J.L.** (2004) Discrimination of suballeles present at the TNFd microsatellite locus using induced heteroduplex análisis. *Genes Immun.* **5**, 76-79.
- Stahl, F.W., Crasemann, J.M. & Stahl, M.M.** (1975) Rec-mediated hotspot activity in bacteriophage λ .III.*Chi* mutations are site-mutations stimulating Rec-mediated recombination. *J. Mol. Biol.* **94**, 203-212.
- Sun, H., Yabuki, A. & Maizels, N.** (2001) A human nuclease specific for G4-DNA. *Proc. Natl. Acad. Sci. USA* **98**, 12444-12449.
- Sutherland, G.R., Baker, E. & Richards, R.I.** (1998) Fragile sites still breaking. *Trends. Genet.* **14**, 501-506.
- Tamaki, K. & Jeffreys, A.J.** (2005) Human tandem repeat sequences in forensic DNA typing. *Leg. Med.* **7**, 244-250.
- Tautz, D. & Schlötterer, C.** (1994) Simple sequences. *Curr. Opin. Genet. Dev.* **4**, 832-837.
- Taylor, J.S., Durkin, J.M. & Breden, F.** (1999) The death of a microsatellite: a phylogenetic perspective on microsatellite interruptions. *Mol. Biol. Evol.* **16**, 567-572.
- Templeton, A.R., Clark, A.G., Weiss, K.M., Nickerson, D.A., Boerwinkle, E. & Sing, C.F.** (2000) Recombinational and mutational hot spots within the human lipoprotein lipase gene. *Am. J. Hum. Genet.* **66**, 69-83.
- Toth, G., Gaspari, Z. & Jurka, J.** (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* **10**, 967-981.
- Treco, D. & Arnheim, N.** (1986) The evolutionary conserved repetitive sequence d(TG-AC)_n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. *Mol. Cell. Biol.* **6**, 3934-3947.
- Turri, M.G., Kuin, C.A. & Porter, A.C.G.** (1995) Characterization of a novel minisatellite that provides multiple splice donor sites in an interferon-induced transcript. *Nucleic Acids Res.* **23**, 1854-1861.

Udupa, S.M., Malhotra, R.S. & Baum, M. (2004) Tightly linked di- and trinucleotide microsatellites do not evolve in complete independence: evidence from linked (TA)_n and (TAA)_n microsatellites of chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **108**, 550–557.

Van Tol, H.H.M., Wu, C.M., Guan, H.C., Ohara, K., Bunzow, J.R., Civelli, O. et al. (1992) Multiple dopamine D4 receptor variants in the human population. *Nature* **358**, 149-152.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G. et al (2001) The sequence of the human genome. *Science* **291**, 1304-1351.

Vergnaud, G. & Denoeud, F. (2000) Minisatellites: mutability and genome architecture. *Genome Res.* **10**, 899-907.

Vogt, P. (1990) Potential genetic functions of tandem repeated DNA blocks in the human genome are based on a highly conserved "chromatin folding code". *Hum. Genet.* **84**, 301-336.

Wahls, W.P. & Moore, P.D. (1998) Recombination hotspot activity of hypervariable minisatellite DNA requires minisatellite DNA binding proteins. *Somat. Cell Mol. Genet.* **24**, 41-51.

Wahls, W.P. & Smith G.R. (1994) A heteromeric protein that binds to a meiotic homologous recombination hotspot: correlation of binding and hot spot activity. *Genes Dev.* **8**, 1693-1702.

Wahls, W.P. (1998) Meiotic recombination hotspots: shaping the genome and insights into hypervariable minisatellite DNA change. *Curr. Top. Dev. Biol.* **37**, 37-75.

Wahls, W.P., Swenson, G. & Moore, P.D. (1991) Two hypervariable minisatellite DNA binding proteins. *Nucleic Acids Res.* **19**, 3269-3274.

Wahls, W.P., Wallace, L. J. & Moore, P.D. (1990) Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. *Cell* **60**, 95-103.

Wahls, W.P. & Moore, P.D. (1990) Homologous recombination enhancement conferred by the Z-DNA motif d(TG)₃₀ is abrogated by simian virus 40T antigen binding to adjacent DNA sequences. *Mol. Cell. Biol.* **10**, 794-800.

Walsh, P.S., Erlich, H.A. & Higuchi, R. (1992) Preferential PCR amplification of alleles: mechanism and solutions. *PCR Methods Appl.* **1**, 241-250.

Weber, J. & Wong, C. (1993) Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**, 1123–1128.

Webster, M., Smith, N.G. & Ellegren, H. (2002) Microsatellite evolution inferred from human-chimpanzee genomic sequence alignments. *Proc. Natl. Acad. Sci. USA* **99**, 8748-8753.

- Weisman-Shomer, P. & Fry, M.** (1994) Stabilization of tetrahelical DNA by quadruplex DNA binding protein QUAD. *Biochem. Biophys. Res. Commun.* **205**, 305-311.
- Weitzmann, N., Woodford, K.J. & Usdin, K.** (1998) The mouse *Ms6-hm* hypervariable microsatellite forms a hairpin and two unusual tetraplexes. *J. Biol. Chem.* **273**, 30742-30749.
- Wierdl, M., Dominska, M., & Petes, T.D.** (1997) Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics* **146**, 769–779.
- Wilder, J. & Hollocher, H.** (2001) Mobile elements and the genesis of microsatellites in dipterans. *Mol. Biol. Evol.* **18**, 384–392.
- Woodford, K.J., Howell, R.M. & Usdin, K.** (1994) A novel K(+)- dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *J. Biol. Chem.* **269**, 27029-27035.
- Woodford, K.J., Weitzmann, N. & Usdin, K.** (1995) The use of K(+)-free buffers eliminates a common cause of premature chain termination in PCR and PCR sequencing. *Nucleic Acids Res.* **23**, 539.
- Xu, X., Peng, M. & Fang, Z.** (2000) The direction of microsatellite mutations is dependent upon allele length. *Nature Genet.* **24**, 396–399.
- Zahler, A.M., Williamson, J.R., Cech, T.R. & Prescott, D.M.** (1991) Inhibition of telomerase by G4-quarter DNA structures. *Nature* **350**, 718-720.
- Zhou, Z., Akgun, E. & Jasin, M.** (2001) Repeat expansion by homologous recombination in the mouse germ line at palindromic sequences. *Proc. Natl. Acad. Sci. USA* **98**, 8326-8333.
- Zhu, Y., Queller, D. & Strassmann, J.E.** (2000) A phylogenetic perspective on sequence evolution in microsatellite loci. *J. Mol. Evol.* **50**, 324-338.

APÉNDICE

Los resultados obtenidos en la presente Tesis Doctoral han sido publicados en las siguientes revistas internacionales (se adjuntan a continuación):

Boán, F., Blanco, M.G., Barros, P., González, A.I. & Gómez-Márquez, J. (2004) Inhibition of DNA synthesis by K⁺-stabilised G-quadruplex promotes allelic preferential amplification. *FEBS Lett.* **571**, 112-118.

Barros, P., Blanco, M.G., Boán, F. & Gómez-Márquez, J. (2005) Heteroduplex analysis of minisatellite variability. *Electrophoresis* **26**, 4304-4309.

